

```

/product="ATP synthase"
/protein_id="AAB87466.1"
/db_xref="GI:2641643"
/translation="MAGKIVQIIICAVIDVEFPDAPVKYDALVESGLTLEVOQOL
GGGLVRCIALGHSIDLKRLKVENGNPQVPGVTKTGRIMNVLGEPIDEKPIGEV"

BASE COUNT      97 a      54 c      79 g      77 t

ORIGIN

Query Match      1.8%; Score 35; DB 1; Length 307;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1308 caaatgcagtaacccaaagtatatgctccttaaa 1342
|||||
Db 58 CAAGATGACGTACCAAACTATATGATGCTTMAA 92

RESULT 6
AF037146      793 bp      DNA      BCT      19-MAY-1998
LOCUS      Salmoneella enterica strain S83769 ATP synthase beta subunit (atpD)
DEFINITION      gene, partial cds.
ACCESSION      AF037146
VERSION      AF037146.1 GI:3138947
KEYWORDS
SOURCE      Salmoneella enterica.
ORGANISM      Salmoneella enterica.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmoneella.
REFERENCE      1 (bases 1 to 793)
AUTHORS      Christensen, H. and Olsen, J.E.
TITLE      Phylogenetic relationships of Salmoneella based on DNA sequence
comparison of atpD encoding the beta subunit of ATP synthase
JOURNAL      FEMS Microbiol. Lett. 161 (1), 89-96 (1998)
MEDLINE      98222522
REFERENCE      2 (bases 1 to 793)
AUTHORS      Christensen, H. and Olsen, J.E.
TITLE      Direct Submission
JOURNAL      Submitted (08-DEC-1997) Veterinary Microbiology, the Royal
Veterinary and Agricultural University, Bulowstvej 13, Copenhagen
1870 Fred. C., Denmark
FEATURES
source      Location/Qualifiers
1. .793
/organism="Salmoneella enterica"
/strain="S83769"
/serotype="T11a 51:z4,z23:-"
/db_xref="taxon:28901"
<1. .>793
/gene="atpD"
<1. .523
/codon_start=2
/gene="atpD"
/transl_table=11
/product="ATP synthase beta subunit"
/protein_id="AAC16682.1"
/db_xref="GI:3138948"
/translation="NGNEKLVLVEVOOOLGGIVRTIAMGSDGLRGLDVKDLHPHIE
VPGKATIGRIMNVLGEPIVDMKGEIPEERMAIHRAPSEELNSOELLEGTIVYD
LMCPFRKGRKVLFGAGVGVKTVNMELIRNIAIHSIGSVAGCERRENDPPIHE
MTDSNVIDKVSILV"
BASE COUNT      181 a      197 c      239 g      176 t

ORIGIN

Query Match      1.6%; Score 32; DB 1; Length 793;
Best Local Similarity 100.0%; Pred. No. 6.8e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1920 atggcgaaataatccgtatgaagtcgtga 1951
|||||
Db 581 ATGGCGGAAAAATTCGATGATGAAGTCGTGA 612

```

```

RESULT 7
AF037147      793 bp      DNA      BCT      19-MAY-1998
LOCUS      Salmoneella enterica strain u24 ATP synthase beta subunit (atpD)
DEFINITION      gene, partial cds.
ACCESSION      AF037147
VERSION      AF037147.1 GI:3138949
KEYWORDS
SOURCE      Salmoneella enterica.
ORGANISM      Salmoneella enterica.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmoneella.
REFERENCE      1 (bases 1 to 793)
AUTHORS      Christensen, H. and Olsen, J.E.
TITLE      Phylogenetic relationships of Salmoneella based on DNA sequence
comparison of atpD encoding the beta subunit of ATP synthase
JOURNAL      FEMS Microbiol. Lett. 161 (1), 89-96 (1998)
MEDLINE      98222522
REFERENCE      2 (bases 1 to 793)
AUTHORS      Christensen, H. and Olsen, J.E.
TITLE      Direct Submission
JOURNAL      Submitted (08-DEC-1997) Veterinary Microbiology, the Royal
Veterinary and Agricultural University, Bulowstvej 13, Copenhagen
1870 Fred. C., Denmark
FEATURES
source      Location/Qualifiers
1. .793
/organism="Salmoneella enterica"
/strain="u24"
/serotype="T11a [Arizonae]"
/db_xref="taxon:28901"
<1. .>793
/gene="atpD"
<1. .523
/codon_start=2
/gene="atpD"
/transl_table=11
/product="ATP synthase beta subunit"
/protein_id="AAC16683.1"
/db_xref="GI:3138950"
/translation="NGNEKLVLVEVOOOLGGIVRTIAMGSDGLRGLDVKDLHPHIE
VPGKATIGRIMNVLGEPIVDMKGEIPEERMAIHRAPSEELNSOELLEGTIVYD
LMCPFRKGRKVLFGAGVGVKTVNMELIRNIAIHSIGSVAGCERRENDPPIHE
MTDSNVIDKVSILVGMNEPGRNRLVALTGLMAEKFEDEKRDVLLFVDNIRYTLA
GTEVSALIGRMSAVGVOPTLAEMGVLOERTISTRTGITSVOAV"
BASE COUNT      182 a      195 c      238 g      178 t

ORIGIN

Query Match      1.6%; Score 32; DB 1; Length 793;
Best Local Similarity 100.0%; Pred. No. 6.8e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1920 atggcgaaataatccgtatgaagtcgtga 1951
|||||
Db 581 ATGGCGGAAAAATTCGATGATGAAGTCGTGA 612

```

TITLE Phylogenetic relationships of *Salmonella* based on DNA sequence comparison of atpD encoding the beta subunit of ATP synthase

JOURNAL FEMS Microbiol. Lett. 161 (1), 89-96 (1998)

MEDLINE 98222522

REFERENCE 2 (bases 1 to 793)

AUTHORS Christensen, H. and Olsen, J. E.

TITLE Direct Submission

JOURNAL Submitted (08-DEC-1997) Veterinary Microbiology, the Royal Veterinary and Agricultural University, Bulowsvej 13, Copenhagen 1870 Fred. C., Denmark

FEATURES

source

1. .793

/organism="Salmonella enterica"

/strain="JEO307"

/serotype="IIIB 61:i:2"

/db_xref="taxon:28901"

<1. .>793

/gene="atpD"

<1. .>793

/gene="atpD"

<1. .>793

/codon_start=2

/transl_table=11

/product="ATP synthase beta subunit"

/protein_id="AAC16684.1"

/db_xref="GI:3138952"

/translation="NGNEKLVLEFVQOOLGGIVRTIAMGSSDGLRGLDYKLEHPLEVPVKATIGRIMNVIGEPVDMKGEIGEEERMAIHRAPSYEELSNSOELLETKIVIDLMPFPARGKGVGLFGAGVGTVMNMELIRNIAIHSIGSVFAGVGEGRNDPFIHMDNSVIDKSVLYVGOMNEPPGNRLRVALTGLTMAEKFRDEGRDVLFFVDNIYRTLLA GTEVSALGRRMPSAVGYOPTLAEMGVLOERITSTKGSITSVQAV"

BASE COUNT 182 a 194 c 238 g 179 t

ORIGIN

Query Match 1.6%; Score 32; DB 1: Length 793;

Best Local Similarity 100.0%; Pred. No. 6.8e-06;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1920 atggcggaataatccgtgatgaagtcgtga 1951

|||||

Db 581 ATGGCGGAAAATTCCTGATGAAGTCGTGA 612

RESULT 9

AF037149 793 bp DNA BCT 19-MAY-1998

LOCUS *Salmonella enterica* strain DS210/89 ATP synthase beta subunit

DEFINITION (atpD) gene, partial cds.

ACCESSION AF037149

VERSION AF037149.1 GI:3138953

KEYWORDS

SOURCE *Salmonella enterica*.

ORGANISM *Salmonella enterica*; Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; *Salmonella*.

REFERENCE 1 (bases 1 to 793)

AUTHORS Christensen, H. and Olsen, J. E.

TITLE Phylogenetic relationships of *Salmonella* based on DNA sequence comparison of atpD encoding the beta subunit of ATP synthase

JOURNAL FEMS Microbiol. Lett. 161 (1), 89-96 (1998)

MEDLINE 98222522

REFERENCE 2 (bases 1 to 793)

AUTHORS Christensen, H. and Olsen, J. E.

TITLE Direct Submission

JOURNAL Submitted (08-DEC-1997) Veterinary Microbiology, the Royal Veterinary and Agricultural University, Bulowsvej 13, Copenhagen 1870 Fred. C., Denmark

FEATURES

source

1. .793

/organism="Salmonella enterica"

/strain="DS210/89"

/serotype="IIIB 48:i:2"

/db_xref="taxon:28901"

gene <1. .>793

CDS /gene="atpD"

<1. .>793

/gene="atpD"

<1. .>793

/codon_start=2

/transl_table=11

/product="ATP synthase beta subunit"

/protein_id="AAC16685.1"

/db_xref="GI:3138954"

/translation="NGNEKLVLEFVQOOLGGIVRTIAMGSSDGLRGLDYKLEHPLEVPVKATIGRIMNVIGEPVDMKGEIGEEERMAIHRAPSYEELSNSOELLETKIVIDLMPFPARGKGVGLFGAGVGTVMNMELIRNIAIHSIGSVFAGVGEGRNDPFIHMDNSVIDKSVLYVGOMNEPPGNRLRVALTGLTMAEKFRDEGRDVLFFVDNIYRTLLA GTEVSALGRRMPSAVGYOPTLAEMGVLOERITSTKGSITSVQAV"

BASE COUNT 183 a 193 c 238 g 179 t

ORIGIN

Query Match 1.6%; Score 32; DB 1: Length 793;

Best Local Similarity 100.0%; Pred. No. 6.8e-06;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1920 atggcggaataatccgtgatgaagtcgtga 1951

|||||

Db 581 ATGGCGGAAAATTCCTGATGAAGTCGTGA 612

RESULT 10

AF037150 793 bp DNA BCT 19-MAY-1998

LOCUS *Salmonella enterica* strain S109671 ATP synthase beta subunit (atpD)

DEFINITION gene, partial cds.

ACCESSION AF037150

VERSION AF037150.1 GI:3138955

KEYWORDS

SOURCE *Salmonella enterica*.

ORGANISM *Salmonella enterica*; Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; *Salmonella*.

REFERENCE 1 (bases 1 to 793)

AUTHORS Christensen, H. and Olsen, J. E.

TITLE Phylogenetic relationships of *Salmonella* based on DNA sequence comparison of atpD encoding the beta subunit of ATP synthase

JOURNAL FEMS Microbiol. Lett. 161 (1), 89-96 (1998)

MEDLINE 98222522

REFERENCE 2 (bases 1 to 793)

AUTHORS Christensen, H. and Olsen, J. E.

TITLE Direct Submission

JOURNAL Submitted (08-DEC-1997) Veterinary Microbiology, the Royal Veterinary and Agricultural University, Bulowsvej 13, Copenhagen 1870 Fred. C., Denmark

FEATURES

source

1. .793

/organism="Salmonella enterica"

/strain="S109671"

/serotype="IIIB 60:i:2"

/db_xref="taxon:28901"

<1. .>793

/gene="atpD"

<1. .>793

/gene="atpD"

<1. .>793

/codon_start=2

/transl_table=11

/product="ATP synthase beta subunit"

/protein_id="AAC16686.1"

/db_xref="GI:3138956"

/translation="NGNEKLVLEFVQOOLGGIVRTIAMGSSDGLRGLDYKLEHPLEVPVKATIGRIMNVIGEPVDMKGEIGEEERMAIHRAPSYEELSNSOELLETKIVIDLMPFPARGKGVGLFGAGVGTVMNMELIRNIAIHSIGSVFAGVGEGRNDPFIHMDNSVIDKSVLYVGOMNEPPGNRLRVALTGLTMAEKFRDEGRDVLFFVDNIYRTLLA GTEVSALGRRMPSAVGYOPTLAEMGVLOERITSTKGSITSVQAV"

BASE COUNT 181 a 193 c 240 g 179 t

ORIGIN

Query Match 1.68; Score 32; DB 1; Length 793;
 Best Local Similarity 100.0%; Pred. No. 6.8e-06;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1920 atggcgaaataatccgtatgaagtcgtga 1951
 ||||||||||||||||||||||||||||||||
 Db 581 ATGGCGAAAAATTCGTCATGAAGCTCGTGA 612

RESULT 11
 LOCUS BACATPA 7307 bp DNA BCT 29-NOV-1995
 DEFINITION B.megaterium ATP synthase i,a,c,b,delta,alpha,gamma,beta and
 epsilon subunit genes, complete cds, and ORF.
 ACCESSION M20255 J04455 M18352 M23924
 VERSION M20255.1 GI:142553
 KEYWORDS ATP synthase; ATP synthase I subunit; ATP synthase alpha subunit;
 epsilon subunit; ATP synthase beta subunit; ATP synthase delta subunit; ATP synthase
 B.megaterium (QM B1551) DNA; clones pWSB100, pCAH1.3, and pWPC208.
 SOURCE Bacillus megaterium
 ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group;
 Bacillus/Staphylococcus group; Bacillus.
 1 (bases 5401 to 7307)
 Hawthorne,C.A. and Brusilow,W.S.
 Sequence of the genes for the beta and epsilon subunits of the ATP
 synthase of *Bacillus megaterium* QM B1551
 J. Bacteriol. 155 (2): 926-931. (1988)
 2 (bases 1 to 5400)
 Brusilow,W.S., Scarpetta,M.A., Hawthorne,C.A. and Clark,W.P.
 Organization and sequence of the genes coding for the
 proton-translocating ATPase of *Bacillus megaterium*
 J. Biol. Chem. 264 (3): 1528-1533 (1989)
 89109162
 Draft entry and computer-readable sequence for [2] kindly submitted
 by W.S.A. Brusilow, 09-NOV-1988.
 FEATURES
 source
 1. 7307
 Location/Qualifiers
 CDS
 /organism="Bacillus megaterium"
 /db_xref="taxon:1404"
 <1..264
 /note="ORF (AA at 1); putative"
 /codon_start=1
 /transl_table=11
 /product="unknown protein"
 /protein_id="AA82518.1"
 /db_xref="GI:1088271"
 /translation="EFGKKILQKGYELMKDRNRHPLQAMALVSIILSQLAGSVLIGL
 FSGRWADMDLDEPLFLILGLGLAAGVFMALRTVHFFSGD"
 271..642
 /codon_start=1
 /transl_table=11
 /product="ATP synthase i subunit"
 /protein_id="AA82519.1"
 /db_xref="GI:142554"
 /translation="MODLOHVEPRRLSYILYLALYVLCMGFTSYKAVFAGLILGTAL
 SLYNLMNLVRKEGFGALDEGKRPISGTVVRFAALAVYITISYPTFHISVV
 GLMTYVYIIDLIVQNMRRR"
 661..1371
 /codon_start=1
 /transl_table=11
 /product="ATP synthase a subunit"
 /protein_id="AA82520.1"
 /db_xref="GI:142555"
 /translation="MGHESPTIEFLGLFPGSGLNLMYVYASIVFLIAVLCTRTLAMK
 PSGAONTEWVLDVFKGLVNSMDKRGREFLTJLVTLMTYFVSNMGLPFAIVIDH
 NLMWKSPTADPAITLTLLAVVYVLSHYGKMRGSAYTKDYFKPMALFLPLKIIIEF
 ANTLTILGLYGNITAGETILSLAGLATTTGTLTGAIAPMLLMQGSIEFGAIOAF
 IFTMLTMYLSHKVSSDH"
 1462..1674

CDS
 /codon_start=1
 /transl_table=11
 /product="ATP synthase c subunit"
 /protein_id="AA82521.1"
 /db_xref="GI:142556"
 /translation="MGLIASAIGIALAGIGNGLIVSKTIEGTARPEARGLTIS
 MMFVGVIAVEALPIIAVIAFMVQCK"
 1777..2295
 /codon_start=1
 /transl_table=11
 /product="ATP synthase b subunit"
 /protein_id="AA82522.1"
 /db_xref="GI:142557"
 /translation="MAVSNMFYGAAGINGDILRQVWFLILALLOKFAFGPWGI
 MKRREHIGEIDEKQNEAKKILVEERELTKSROQVQWMEKARSAQKKEI
 VAAARESERLKAANKQETEDQKDAVALRQVAVSLVLSAKVIELKSDQDEKL
 IHEYIOEVDGR"
 2292..2837
 /codon_start=1
 /transl_table=11
 /product="ATP synthase delta subunit"
 /protein_id="AA82523.1"
 /db_xref="GI:142558"
 /translation="MSQPAVKARVALALPOLATEKOMIDEMODLOIIVEFPAKPEL
 MVLTHPKTIERKQFSEAFARLSELPVQVHLLLEHRIQIYSEWKEVRLANE
 VAGVADATYVSKPLSADKRAISQSFASKVKHNLINISYDKYIGVKLRIGNRI
 YGSSISKLETHRGHLAHS"
 2853..4361
 /codon_start=1
 /transl_table=11
 /product="ATP synthase alpha subunit"
 /protein_id="AA82524.1"
 /db_xref="GI:142559"
 /translation="MSIKAEISALIKQOIEYQSEIKYSDVGVYIOWDGIARAHL
 DNWAGELVEFSGWGAQNLKENNVGIIIGPYTEIEGDEVRRTGIMVPGEQ
 LIGRVYNSGQVVDGIGPEYETKTRPIEGAGVMDRKSVHPELOTGKAIDALVPIG
 RGORELIGDROTQTSVAIDPILNOKDMQCIYVAGOKESYRNVETLRKGL
 DTVIVTASASQAPLLPLAPYAGVYMGEEFNKGHVYITIDLTQKASATRELSL
 LRRPGREAYPGDVFYILHSRLLEPERAKISDAGGSLVAPLETQAGVSAVITPNV
 ISITDGOITFQSDLEFFSGVRAINAGLSVRGSGQIKAMKVGATGLDASREL
 ESFAQFGSDLDQATQAKLNRGARVETILKQGLHKLPRVQVAVYALTKGLDQVY
 SDITRPEDEYLTWLESNKREVEISIRTTGGLPEAGLFEALDEFKTFIASH"
 4467..5324
 /codon_start=1
 /transl_table=11
 /product="ATP synthase gamma subunit"
 /protein_id="AA82525.1"
 /db_xref="GI:142560"
 /translation="NASLSDIOTRITSTFKTSQITKAMMVSAAKLNRQAKSFV
 IWEKIOEVVSSVALSGASHBPLTARSKTKGYIVTSDRLAAYNSNLIKRSQA
 IERHQSDEYVIAIGRVGRDFPVKRGIPVLETTGGLADQVADIGIASQTVOMF
 ADGTFDELYLYNHFINTISQVETEKKLPLDLIDGLVQVEEPESEELTEVLLP
 OYAESLIYGLLDGKASEHAHAWTAMK SATDNKDLINLITLSYNARQAATQETTE
 IVGAAALE"
 5361..6779
 /codon_start=1
 /transl_table=11
 /product="ATP synthase beta subunit"
 /protein_id="AA82526.1"
 /db_xref="GI:142561"
 /translation="MFKGVYQIMGPYVDYKRDNGHLPATYNAKLSHKPSASVAYI
 ELTLEVAHLGDNIVRYVAMSSTGDLVRLEVEDGALISVPGVDTLGRVNVJGEK
 IDLDAPIDAGARRDPIHRQAPFENLSQAEILTEGIRKVDLLAAYIKGKIGLFGGA
 GVGTIVLLOELINNTAIEHGGISVAGVETRENDYIHEMTDSGVIKTKYAMFGGA
 NEPPGARQORVATGLTMAEYFRDEQGDVLEFDINIFRTOAGSEVALGTMPSAGM
 YOPTLATEMGLOERITSTSVGSVSIAIYVADYDPPAPATFAHLDATYINERK
 LSEKTIYPAVDPLASTSRALSPEIYGEHEIARAQVOOTLOCKYKLODIATILGDEL
 SDEKLVVPRARKVOFFLSQTSVAQFTGQKSYIVAKETVAGFEIILEGKIDHLPED
 ARLRLGLEEVLENKRGVEV"
 6806..7210
 /codon_start=1
 /transl_table=11
 /product="ATP synthase epsilon subunit"

BASE COUNT 2274 a 1331 c 1639 g 2063 t
 1 bp upstream of EcorI site.

Query Match 1.38; Score 26; DB 2; Length 7307;
 Best Local Similarity 100.0%; Pred. No. 0.015;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1779 gtagctagcgtacgcgtgaagttaa 1804
 ||||||||||||||||||||||||||||
 DB 5922 GTAGGTAGCGGTACGCTGAAGTAA 5947

RESULT 12
 LOCUS AF008210/c
 DEFINITION AF008210 34699 bp DNA BCT 31-JAN-1998
 Buchnera aphidicola genomic fragment containing (chaperone Hsp60)
 groEL, DNA biosynthesis initiating protein (dnaa), ATP operon
 (agpGDAHEB), and putative chromosome replication protein (gidA)
 genes, complete cds; and termination factor Rho (rho) gene, partial
 cds.

ACCESSION AF008210
 VERSION AF008210.1 GI:2827005
 KEYWORDS Buchnera aphidicola.
 SOURCE Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 ORGANISM Buchnera aphidicola
 1 (bases 1 to 34699)
 Clark, M.A., Baumann, L. and Baumann, P.
 Sequence analysis of a 34.7-kb DNA segment from the genome of
 Buchnera aphidicola (endosymbiont of aphids) containing groEL,
 dnaa, the atp operon, gidA, and rho
 Curr. Microbiol. 36 (3), 158-163 (1998)
 98184963

REFERENCE JOURNAL
 MEDLINE 2 (bases 1 to 34699)
 Clark, M.A., Baumann, L. and Baumann, P.
 Direct Submission
 Submitted (13-JUN-1997) Microbiology Section, Univ. of Calif. at
 Davis, Davis, CA 95616-8665, USA

FEATURES
 SOURCE Location/Qualifiers
 1. 34699
 /organism="Buchnera aphidicola"
 /specific_host="Schizaphis graminum"
 /db_xref="taxon:9"
 <1.437
 /codon_start=3
 /transl_table=1
 /product="39-kDa hypothetical protein"
 /protein_id="AAC38098.1"
 /db_xref="GI:2827006"
 /translation="MHTRLPIVPIPNRTSDLCQIFSNISVLIIVTHINHPQELNEOL
 SDSLKLTKKSNVILNQSVLKININDNAIIIAELSSRCENNIIPYLIHLIDKVGTS
 HEIVSNKAKSIIDLMKMTSGFLVPLRVFNGSKDNKLII"
 complement(558..2213)
 /gene="groEL"
 complement(558..2213)
 /gene="groEL"
 /codon_start=1
 /transl_table=1
 /product="chaperone Hsp60"
 /protein_id="AAC38099.1"
 /db_xref="GI:2827007"
 /translation="MSKMGADVKGFGNEARIKMLRGVNLADAVKVTLGPGRRNVLD
 KSPGAPSTTKGVSVAREIELEDENFGNAQGVKAVSKANDAGDGTITATLLAQSI
 VNEGLKAAAGMPPMDLRGIDKAVISAVEELKNLSVPCSDSKAITOGTISANDK
 VGLALAEAMERYGNDGVITVEEGRLQNELELYVGKMPDRGCLSPYFINKEGTIVEL
 ENPYILMDKRIISNREMLPILESVAKGKPLIISDLSEBALATLVNPMRGIVKY
 AAVAPGFGDRRKMLDIDISITLGGSVISEELMLEKSTLEDLGQAKRVVINKDTT"

gene
 CDS

ITGSGSKOAIOSRIGOIROEIOEATSDPKKEKLEIRLAKISGVAVLKVGATEVEN
 KKKARVBDALHATRAAVEGVAGGVALVPAKISLNRGHNDDQVIGVALRAM
 EAPLQIYNSGSEBSVYTNKKDKKGNGVYAAADNDEGMDIDFGIDPTKYTRBALQ
 VAAVAGMLTTECEVYTDLPREDKSSDVAA SPAGMGNGMGAM"
 complement(2262..2552)
 /gene="groES"
 complement(2262..2552)
 /gene="groES"
 complement(2262..2552)
 /gene="groES"
 /codon_start=1
 /transl_table=11
 /product="chaperone Hsp10"
 /protein_id="AAC38100.1"
 /db_xref="GI:2827008"
 /translation="MKIRPLHVRVLYKRNELKSGAGIVLGSAAKSTRGITAVG
 NGRVLDNGOIKPLDVGGVYVFNPEGYGAKEIKINELLINSDILALVE"
 complement(2795..2870)
 /product="trna-phe"
 complement(2884..4254)
 /gene="trna-phe"
 complement(2884..4254)
 /gene="trna-phe"
 /codon_start=1
 /transl_table=11
 /product="trna-methyltransferase"
 /protein_id="AAC38101.1"
 /db_xref="GI:2827009"
 /translation="MIRNDITIAQVYCPKRSANGILRVSGIHANOVAFVAVICKPKPR
 FAYSKFEFSKRVLDGSLWFPAPSLDEVDLELOGHGPFIMDLIRLCLCKN
 IKIRIAPQGEFCORAFNLGKIDILQAEADLINSLETESVYRSLNSLHGNEFFYLOK
 IIRKLEEFRTIASIDFSEENIDPDFNIFPSNPEKLNDFIKINIVSEGLIEBA
 KRIYVIGPNAKGSLLNLVNCSDRAIVLDELPTGVDLYENINIGISCEIYDAGL
 RETEDKIERIGIORSMEIKNSDHYLYMDKTSISLDOKKTSLOPKOISSYNIQVY
 RVLNKNDLVDFEGITKIEMLTISALTGQSIDLKRHSRQKDSOEGFIIARR
 RHIIHIDLSYCELDLAKQKMLKYNIELLESINTINKLIGETTSDDLKRIFS
 TFCIGK"
 complement(4297..5907)
 /function="unknown"
 /codon_start=1
 /transl_table=11
 /product="60-kDa protein"
 /protein_id="AAC38102.1"
 /db_xref="GI:2827010"
 /translation="MELORNFPIFAFLVFSFLLQMAWQSOLKKNKNEEFSFHLN
 HKREDKQIITKNDVLRVLVNMVGGDIEEASLDYKQLNSPENLKLLEKSPVVOA
 QSGILGKDGDNSTNKTRPLYSAKKNFEELGHNEKELRVPTTAKNGITVYFIK
 PGKYNVEYEDINNLSNKKLELNFGLKOTVLPKRDYISGFALQTRGAAVSS
 DAKYKKFPDNIANHEHLITTHGWTAMQOYFVAVAVDPDNISSNIIYTSILDND
 GIATIGYKSSLINIPSNRYIIRKSKMIGPEKQOEMALVAPNDLTVDFGLWFLSOP
 LFKLSTIHNFIGNWFSIILIPFMAIATPYTKAOTSMKRELOPIMELKKNF
 GHKQKRSKEMALYKKEKINPLGGCLPVIQMPFSLYMLIGSVELRHAPFLFWI
 KDLSDDPYVYLPFPGELTFEFTORTSSNNISDPFOCKIMHMFIFTVFLEPPSGL
 VLYIYNSLVITIIQKYLISNFKKQKR"
 complement(5910..6167)
 /codon_start=1
 /transl_table=11
 /product="10-kDa hypothetical protein"
 /protein_id="AAC38103.1"
 /db_xref="GI:2827011"
 /translation="MVKLSTIVFCLTFEVSITYONYISFEMPSRCRYPTCSITYMIS
 LRKFGVIGKILITLIRLFKCHPHOGGEDVPLKIKDSEY"
 complement(6146..6490)
 /gene="trna-p"
 complement(6146..6490)
 /gene="trna-p"
 /codon_start=1
 /transl_table=11
 /product="RNase P protein component"
 /protein_id="AAC38104.1"
 /db_xref="GI:2827012"
 /translation="MLNFTFKKSKLLKSTNFQYVFSNPCKNTPFHINILGRSMLGH
 PRLGISIRKNIKHAAYRRNKIKRLIRETFRLQRLISMDEVIAKKNIYVLLNKKIV
 NILEYISNVOR"
 complement(6504..6647)

gene
 CDS


```
gene
CDS
/gene="rpmh"
/complement(6504..6647)
/gene="rpmh"
/codon_start=1
/transl_table=11
/product="50S ribosomal subunit protein L34"
/protein_id="AAC38105.1"
/db_xref="GI:2827013"
/transl_table="MKRTFQPSILKRNNSHGFFRMTATKNGRYILSRRAKLRTPLTV
SSK"
7066..8430
/gene="dnaa"
7066..8430
/gene="dnaa"
/codon_start=1
/transl_table=11
/product="DNA biosynthesis initiating protein"
/protein_id="AAC38106.1"
/db_xref="GI:2827014"
/transl_table="MSICIMKOCCLDRLOSELSTFESFWIRSLKAKLNNNILEIYAPN
QPIIDWYADKYLHFKTLQDFGSGNSPFIFKYYQTSKERRFKNNILQKIQNLNAPK
IMDKIPIFKSSHSNINIKKHSFENFIEGKSNOLRAASQVAKNPGNSYNPLFLYGG
TGLCKTLLHAIIGGILAYKYNKIIYHSHSERFQDWAYKALQNNNAIEKFLYRSVDA
LLIIDIOFPAKERSOEFEFTFALLEGNOQIILTSDRYPKETNGVEDRLKSRGAG
LTVADPELEFTRVAIIKIKADENNIVLSDFAFIKALRSNVBELEGALNRYVNA
NFTHRSTVEFREALDILALQEKIYTIANIQKTVAYYIKIVADLLSRRSRVAR
PRQAMAMAKELIHNLSPEIGDAFSGDHTIVLHACRKIEQIKRENHDIKEDFSMLIR
TLVY"
8435..9535
/gene="dnaN"
8435..9535
/gene="dnaN"
/codon_start=1
/transl_table=11
/product="DNA polymerase III beta subunit"
/protein_id="AAC38107.1"
/db_xref="GI:2827015"
/transl_table="MKFTIIONDIITKNNKKITRYLVKNSIPFLENTLLIOVEDGTLST
TTTNLEIETLISKEIETIRYIIGKTTISGRKTLINICRTISEKRTKMLKNNKMYISSE
NSNTILSTLSADTFPNHONFDYISFKDISNILEMEIKETFSWGDVRYVNLGMLL
EKDKFLSVATDGYRLAISYQDKDINPFSIIIPKAVMELKLNTQPOLNILI
GNSIRIYTKNLIPTTQLEGEYDYKSVLFKEKKNPITINSILKSLRLVAIIAHE
KFCGEIKENGKFKVLSNDQEEETAEDELFEIDYGEKIEISINVIYLDVIYNNKSE
NIALFLNKSXSIIQEAENSSNAVVMALKR"
9690..12101
/gene="gyrB"
9690..12101
/gene="gyrB"
/codon_start=1
/transl_table=11
/product="DNA gyrase subunit B"
/protein_id="AAC38108.1"
/db_xref="GI:2827016"
/transl_table="MIDYDSSKIKILRGILDVARRKPMYIGTDGDSGLHHMVEIY
DNSIDEALAGCFKEIKVYIHSNDSYKIDQDREGIPTDIHPEKISAAEIVMTVLHSGG
KEDNASYKISGGLGCVGISVYNALSEKELIYKKYKRYROOYENGNEPSPCLIGAT
EITCTYIRFPSTYKIFTRKTOFOYEILSKRLRELSFNSNTLIHLKDRINDEKTYH
KGGIKAFKFLNKKKPTNLNIEFYRCIKNOIELEIALQNNNSHQENILCFTNMPK
DGGTHLSGFRAGITRTLMLEIRGYNKKNTIITGEDTREGTAIIISIKIPDFSS
QTKRLVSEVYSIESLINENLIEYLENDIAKFIQKTIINSAKIREAARRAEIN
KKKGLDGLTLPGLSDQENDPKISEIYELVSDGAGSAGKAGRNRRKNOAILPLKGI
LNVGKTEPKDIIISOLETSLTALGCSITRSEYSDKIRYNHITIMPDADYGAHIRT
LLLFYRQDELIRKGVYTAOPPLIKVKKGKDEKTIKNDENMKYQIKALKEIVI
KNNKNECNKTIKEFOKIVSDFNHIIQIKKKKNKYPPLINELIYHRLYNLKNENY
VQNNIEKLVENLKNKNNNTIYTSKIKENDSDIFEPSIKLSRYANHQYDLDKNDFE
SKEVFLITNLEGEKFKFOENENFIEKGEIKYKINDIKNTLEMLIKETKRGFEVORYG
LGENPQDLNMTYNNPETRNMLQYTIKDAISANNLFTLMGDIVPRRRKFIQKNALKA
ENIDI"
complement(12206..12622)
/gene="atpC"
/complement(12206..12622)
/gene="atpC"
/codon_start=1
```

```

/transl_table=11
Query Match 1.3%; Score 25; DB 1; Length 34699;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 424 actaaagcgatggaatggttcgtg 448
|||||
Db 14881 ACTAAAGCGATGGAATGGTTCGTG 14857
```

Search completed: October 28, 2001, 17:56:17
Job time: 8411 sec

Db 361 aaaaagcagtgctaaagagataagaaccaaaatcgagtgtaaaagtlacaaaaa 420
QY 421 attractaaaacgagtggaatggtgtgctgcctcgaaaatgctgaagaacgacgcatg 480
Db 421 attactaaagcagatggaatggtgtgctgcctcgaaaatgctgaagaacgacgcatg 480
QY 481 tctcttcaacgccccttaatcgaagaacaatacgtlaacgltgattagccagtttccaaagca 540
Db 481 tctcttcaacgccccttaatcgaagaacaatacgtlaacgltgattagccagtttccaaagca 540
QY 541 acgactggttacaaacgcatccttcttagtggtgcggaagtaaaaaagtggtgacatg 600
Db 541 acgactggttacaaacgcatccttcttagtggtgcggaagtaaaaaagtggtgacatg 600
QY 601 gttgtgcacacagatcggtgctctgtgtgtgcttaaacggtgaacttggttaaaacgta 660
Db 601 gttgtgcacacagatcggtgctctgtgtgtgcttaaacggtgaacttggttaaaacgta 660
QY 661 ttaaatgaaatgaaagaatggaagaagaagaatgttccgltcaattgagtttaacggt 720
Db 661 ttaaatgaaatgaaagaatggaagaagaagaatgttccgltcaattgagtttaacggt 720
QY 721 tctaatctcaacattttccaatctttgggaattaaatttaacccaagatctcaggt 780
Db 721 tctaatctcaacattttccaatctttgggaattaaatttaacccaagatctcaggt 780
QY 781 attgtgtactccctcgtctgtgacagttlaattgttcacgtcaattctatgattgct 840
Db 781 attgtgtactccctcgtctgtgacagttlaattgttcacgtcaattctatgattgct 840
QY 841 tataaaaaaggggaagtagatgctgtgctgatttagttatacaaaatttataacagatg 900
Db 841 tataaaaaaggggaagtagatgctgtgctgatttagttatacaaaatttataacagatg 900
QY 901 tcgcaaaaagccagatgttgaaaataatltccattacagaattagaaatgagatgaatta 960
Db 901 tcgcaaaaagccagatgttgaaaataatltccattacagaattagaaatgagatgaatta 960
QY 961 ggcgaagaagaacaagttctggatataattacgaacctgtgacgaagttatttagat 1020
Db 961 ggcgaagaagaacaagttctggatataattacgaacctgtgacgaagttatttagat 1020
QY 1021 aatttattggttcgttatttgaatctcaggtttatacgaacgacgcttgaaaacctgct 1080
Db 1021 aatttattggttcgttatttgaatctcaggtttatacgaacgacgcttgaaaacctgct 1080
QY 1081 tctggaagcagccgctcgaaatgctgcgcataagaaagcagaacagataacggttaacta 1140
Db 1081 tctggaagcagccgctcgaaatgctgcgcataagaaagcagaacagataacggttaacta 1140
QY 1141 attaatgagttacagttagctatacaaaagctcgtaacgaagtattacaatgaaatta 1200
Db 1141 attaatgagttacagttagctatacaaaagctcgtaacgaagtattacaatgaaatta 1200
QY 1201 aatgaaattgttgcgcggtgcgcagcagaatttaacaataagagatcggttaagtgaactg 1260
Db 1201 aatgaaattgttgcgcggtgcgcagcagaatttaacaataagagatcggttaagtgaactg 1260
QY 1261 gaaaattgtaacaatcactcgtgtggttattgaaatgaaatccacacaagtgacgtac 1320
Db 1261 gaaaattgtaacaatcactcgtgtggttattgaaatgaaatccacacaagtgacgtac 1320
QY 1321 caaaaagttatagtgcttaaaatgttgaacaaggtttagttacttgaagtccaacaacat 1380
Db 1321 caaaaagttatagtgcttaaaatgttgaacaaggtttagttacttgaagtccaacaacat 1380
QY 1381 taggt 1440
Db 1381 taggt 1440
QY 1441 taagcgttaacaatacgaataacccaatttctgttcaagtgtggaacgaataacattgggtc 1500
Db 1441 taagcgttaacaatacgaataacccaatttctgttcaagtgtggaacgaataacattgggtc 1500

QY 1501 gttacatgaacgattatggttgaacccaatcgaatgacgaagttgaatccgtgtcagaagaga 1560
Db 1501 gttacatgaacgattatggttgaacccaatcgaatgacgaagttgaatccgtgtcagaagaga 1560
QY 1561 attggttattacacggtgcgcgcacccaagttatgaagaacaatacagatgaacttt 1620
Db 1561 attggttattacacggtgcgcgcacccaagttatgaagaacaatacagatgaacttt 1620
QY 1621 tagaagcgggaattaaagttacatcagtttagtttgcctgttgcgaaggggtlaaagtag 1680
Db 1621 tagaagcgggaattaaagttacatcagtttagtttgcctgttgcgaaggggtlaaagtag 1680
QY 1681 gtttattcgt 1740
Db 1681 gtttattcgt 1740
QY 1741 tcgcaattggaacactcaggttactcgtcttctgtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1800
Db 1741 tcgcaattggaacactcaggttactcgtcttctgtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1800
QY 1801 gtaacgactctatcatgagatgaaagacttaacgtattagataaagtgctctgttt 1860
Db 1801 gtaacgactctatcatgagatgaaagacttaacgtattagataaagtgctctgttt 1860
QY 1861 atggtcaaatgaaacgagccaccaggttaacggtttaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1920
Db 1861 atggtcaaatgaaacgagccaccaggttaacggtttaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1920
QY 1921 tggcgggaaaaattccgtgtgaaaggt 1972
Db 1921 tggcgggaaaaattccgtgtgaaaggt 1972

Search completed: October 28, 2001, 17:58:55
Job time: 4474 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 15:34:11 ; Search time 1454.83 Seconds
(without alignments)
12813.200 Million cell updates/sec

Title: US-09-545-199c-3
Perfect score: 1972
Sequence: 1 agcgggcattgctcagc.....gtcttattcttgatataa 1972

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 10228115 segs, 4726426750 residues

Word size : 25

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

```

```

190: gb_est110:*
191: gb_est111:*
192: gb_est112:*
193: gb_est113:*
194: gb_est114:*
195: gb_est115:*
196: gb_est116:*
197: gb_est117:*
198: gb_est118:*
199: gb_est119:*
200: gb_est120:*
201: gb_est121:*
202: gb_est122:*
203: gb_est123:*
204: gb_est124:*
205: gb_est125:*
206: gb_est126:*
207: gb_est127:*
208: gb_est128:*
209: gb_est129:*
210: gb_est130:*
211: gb_est131:*
212: gb_est132:*
213: gb_est133:*
214: gb_est134:*
215: gb_est135:*
216: gb_est136:*
217: gb_est137:*
218: gb_est138:*
219: gb_est139:*
220: gb_est140:*
221: gb_est141:*
222: gb_est142:*
223: gb_est143:*
224: gb_est144:*
225: gb_est145:*
226: gb_est146:*
227: gb_est147:*
228: gb_est148:*
229: gb_est149:*
230: gb_est150:*
231: gb_est151:*
232: gb_est152:*
233: gb_est153:*
234: gb_est154:*
235: gb_est155:*
236: gb_est156:*
237: gb_est157:*
238: gb_est158:*
239: gb_est159:*
240: gb_est160:*
241: gb_est161:*
242: gb_est162:*
243: gb_est163:*
244: gb_est164:*
245: gb_est165:*
246: gb_est166:*
247: gb_est167:*
248: gb_est168:*
249: gb_est169:*
250: gb_est170:*
251: gb_est171:*
252: gb_est172:*
253: gb_est173:*
254: gb_est174:*
255: gb_est175:*
256: gb_est176:*
257: gb_est177:*
258: gb_est178:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description
---------------	----------------	-----------------	----------	-------------

No matches found

Search completed: October 28, 2001, 17:10:56
Job time: 5805 sec

THIS PAGE BLANK (USPTO)

FEATURES
source
1. 13828
/organism="Haemophilus influenzae Rd"
/db_xref="taxon:71421"
55. .1168
/gene="HI0470"
55. .1168
/note="similar to SP:P06986 GB:002071 GB:X03416 PID:41695
PID:41710 percent identity: 60.40; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="histidinol-phosphate aminotransferase (hisc)"
/protein_id="AAC22129.1"
/db_xref="GI:1573449"
/translation="MNDYNSHILQMKKTIMTITLSRONIQALTPYQARKLGNGTI
WLNANEYPTSPFQSLKDNRYPEPQPVQAVANAGVSTENLVITRGDEGIEL
IHTECEPKODALFCPTPYGMVAVAETAGVLTPELDFOLNLEIKHNDYK
VYVCSNNPPTGNILKQSDIIDLQITAGKATVYDEATECEPPEASVNLKNTPHL
AIIPTLSKAFALAGLRGCFVLANPELIDLSKVIAPIPIVPSADIAEQALPSNAT
VQALTELQNSKQMLAKALVLHQVEKYESSANLLIKQNGQAVFKALMBQGIILR
DONKTLHLQNCIRITVGTRECEKVEAIKEVK"
1280. .2368
/gene="HI0471"
1280. .2368
/note="similar to SP:P06987 GB:X03416 PID:41696 PID:41711
GB:000096 percent identity: 64.97; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="imidazoleglycerol-phosphate dehydratase /
histidinol-phosphatase (hisb)"
/protein_id="AAC22130.1"
/db_xref="GI:1573450"
/translation="MQPTLEIDRGLIDEKPTDFQIDSLKLEKIPVYIPALRLKA
KRYEIVSNQDGLGDAFPQNDPDKPHNVMALPESGITEPDEVLCIPKPRENLCL
KPKTILDIHYIKFNLPIDRSRVISDREDDVOLAENLRATQYDPOKMMDLIAEKL
LGEITVNCGRPRFAEVIROTKEVDIVQVLAQAGVNEIKTVGFEDHMLDQIATH
GCFRNAYOCKGDLWIDENHTVBDTALALQALQAVGDKRGIRAGFVLPDCECAEC
ALDLSGRPMWIKNACFKDKVGFESTELTEHFOSLAFSMLATLHLNVTGNDDHKIE
SLKFAGRTLROAIRIEGNEMPSKGLV"
2434. .3033
/gene="HI0472"
2434. .3033
/note="similar to SP:P10375 PID:41712 GB:000096
PID:1736701 PID:1736710 percent identity: 55.90;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="amidotransferase (hisb)"
/protein_id="AAC22131.1"
/db_xref="GI:1573451"
/translation="MINTIITDGCANLSVAFADRLGYNEITFEDLNKIKSADKLI
LPGVGTANAYNLOEROLIEIOMLTQPVLCICGMQMLTFESEGNTPTNLISCK
TNRIPTDGLPDMQMNRYQVYKNCPLTDGIVONSHYFVHSYAVSPHEHSVAISNG
VNFSAIAKENYGVQFHPERSGKNGLLNKFKVPE"
3069. .3818
/gene="HI0473"
3069. .3818
/note="similar to SP:P10371 PID:41713 GB:000096
PID:1736702 PID:1736711 percent identity: 62.86;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="phosphoribosylformimino-5-aminoimidazole
carboxamide ribotide isomerase (hisA)"
/protein_id="AAC22132.1"
/db_xref="GI:1573452"
/translation="MKOSIITIPALDLINGQVRLHQGVAKQTYSDNPICKQFQYVR

gene
CDS
3800. .4576
/gene="HI0474"
3800. .4576
/note="similar to SP:P10373 PID:41714 PID:598474 GB:000096
PID:1736704 percent identity: 82.03; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="hisF cyclase (hisf)"
/protein_id="AAC22133.1"
/db_xref="GI:1573453"
/translation="MAKRRIIPCLDVRDQVYGVQFNRHEITIGDIVPLAQRVQEGA
DELVAEDYIASDGRVDSWERYIAOVADIPFCVAGIKTTEDEKLEAFAGADKISI
NSPALADPTLISRLADRGVQVIVGIDSMFQDCKYVWVNTGDETRTQTHQL
DMYKEVQCGAGEYILNMNDGLRNGDIDLAQIKYRGVCRPLIASGAGAKMVFED
AFIEAKVDALASVPHKQIIEIGELKSYVQSAIEINSE"
4576. .5241
/gene="HI0475"
4576. .5241
/note="similar to SP:P06989 GB:D43637 GB:002072 GB:X03974
PID:41700 percent identity: 61.22; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="phosphoribosyl-AMP cyclohydrolase /
phosphoribosyl-ATP pyrophosphohydrolase (hisIE)"
/protein_id="AAC22134.1"
/db_xref="GI:1573454"
/translation="MNITKIDQKVNGLPVIQNIISREVLMGLYNEEALTKTKE
RKVFESRTRKORLMTGELSGNFIENVEESLDCDDTLLIWDIPGACHTGNSCFH
OETSPQSEKQOFPANMAWFIKLROHIEKKRADSNSYVTLNIAKSGKKAQKQGE
GVETPLAANQDAKAEVISEATDLYVHLVLLHNDLQMTETIAKIQENHOGIGLHPEG
GNK"
5323. .5475
/gene="HI0476"
5323. .5475
/note="hypothetical protein; identified by Genemark:
putative"
/codon_start=1
/transl_table=11
/product="H. influenzae predicted coding region HI0476"
/protein_id="AAC22144.1"
/db_xref="GI:1573464"
/translation="WIKCIDKQQLNGNIIIFLLKQOYSKEDSKKFTIYFQTVN
YTIQLS"
5505. .6707
/gene="HI0477"
5505. .6707
/note="similar to GB:M23240 SP:P18199 PID:148089 GB:000096
PID:1736569 percent identity: 47.24; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="tyrosine-specific transport protein (tyrP)"
/protein_id="AAC22135.1"
/db_xref="GI:1573455"
/translation="MNKTVGSLTLVAGTMIGAGMLDPLTSAGIGFGLVLLGLMA
LTFESALLFVELYQTAESDAGIGLAEQVFGTGRITAVAILFVLLALIAVYISGGG
SLKDLPEPSGDKGVLLFVYIGSFVIGHSVDKINRVAILFVMLAFAVAVSIML
PEIKRDNMAPPIDKALITISAPVEFTAFGRHSIPSKNKYDGNVKAIRRSIIVGSA
ITLCAYITIMQSLTHLQNEPQILKEDATLNGKATPRAITGSNVAASAVKLFSTL
ALITSELVGLGLECIEDLKRSPNVTAGRISTGLTFIPPLVPALEPFGFIIALG
YAGQFAFYAVAVLPVSLVWKAARRAHANLPYKWMGMLTLIVLVGLVLTISIPAIRA
GLPFPVVG"
complement(6893. .7321)

Source	Organism	Reference Authors Title	Journal
source	Pasteurella multocida.	Fuller, T.E., Kennedy, M.J. and Lowery, D.E.	Journal of Clinical Microbiology
gene	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Pasteurella.	1 (bases 1 to 870)	
CDS	Identification of Pasteurella multocida virulence genes in a septicemic mouse model using signature-tagged mutagenesis	2 (bases 1 to 870)	
	Unpublished	3 (bases 1 to 870)	
	Submitted (24 FEB 2000) to Discovery Research, Pharmacia & Upjohn Animal Health, 7923-25-434, 7000 Portage Road, Kalamazoo, MI 49001-0199, USA	4 (bases 1 to 870)	
	Location/Qualifiers	5 (bases 1 to 870)	
	/organism="Pasteurella multocida"	6 (bases 1 to 870)	
	/db_xref="taxon:747"	7 (bases 1 to 870)	
	/gene="atpC"	8 (bases 1 to 870)	
	/gene="atpG"	9 (bases 1 to 870)	
	/codon_start=1	10 (bases 1 to 870)	
	/translation=1	11 (bases 1 to 870)	
	/product="ATP synthase F1 gamma chain"	12 (bases 1 to 870)	
	/protein_id="AA68408.1"	13 (bases 1 to 870)	
	/db_xref="GI:7716505"	14 (bases 1 to 870)	
	/translation="MACAKEIRFKIASVKSNOITKAMEVYASKMRTOERMSNRPRYSERIVSHVSKATGATGYKRPHPLEDEVKYGVIVSTPGRLGQNLVFKVLENMKEKEDVYVQSLISKSIINFPQSLIKILTODSSIGIDPSYEQILGVSNSIDANMKKEVDVYLVNKFINTMSQKPYLELIPLELDNDELGRKQWVYIYEPDKVLLDNLELVRLSQVYQAAVENLASDEQAAVMVAMKATDANGLINELQLVNRAQASITVNELENYVAGAAIT"	15 (bases 1 to 870)	
BASE COUNT	305 a 132 c 177 g 256 t		
ORIGIN			
Query Match	44.1%; Score 870; DB 2; Length 870;		
Best Local Similarity	100.0%; Pred. No. 2.8e-160;		
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Y	364 atggcaggtgcttaagagataagaacccaacatcgcgaggtgtaaaagtacacaaaatt	423	
Db	1 ATGCGAGGTGCTTAAGAGATAGAACCAAAATCGCGAGTAAAGATACACAAAATTT	60	
Y	424 actaagcagatgaatagt	483	
Db	61 ACTAAAGGAGTAAAGATGTTGTGTCCTCGAATAATGCTTAAAGCGCAAGACGATGCT	120	
Y	484 tcttcacgaccttattcgaacaatactagtaagttatagccacgtttccaaagcacg	543	
Db	121 TCTTCACGCCCTTTATCGAACAACATACGTACGATTAACCGACGTTCCAAACGACG	180	
Y	544 atgtgttaacagatccatcttttagtggatcgcgaaagtaaaaaaagtggcagatgtgt	603	
Db	181 ATTGGTTACAGCATTCATTTTGTAGTGATCGCGAAGTAAAAAAAGTGCGCATGTTGT	240	
Y	604 gtgtccacagatgt	663	
Db	241 GTGTCCACAGATGTCGTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	300	
Y	664 aatgaatgaagaatggaagaagaagaatgttccgttcaatttgagttaatcgtgtct	723	
Db	301 AATGAATGAAGAATGGAAGAAGAAGAATGTTTCCGTTCATTCAGTTTAATGCGTTCT	360	
Y	724 aaatctatcaacttttccaaactttgggaattaaaatttaaccacgaatcagatgt	783	
Db	361 AAATCTATCAACTTTTCCAACTTTTGGGAATTTAAATTTTAACCCAGATTCAGGTAT	420	
Y	784 ggtaatactccctgtttggacagtaattgttttcagtcacattctatgatgtgattat	843	
Db	421 GGTAATCTCCCTGTGTGACAGTTAATGTGTTACGCAATTCATGATGATGATGCTTAT	480	

QY	844	aaaaaagggaagtgaagttgtagttagttataaacaatttataacaagtcg	903
Db	481	AAAAAGGGGAAGTAGTAGTCTGTGATTTAGTTTATACAAATTTATTAACGATGTCG	540
QY	904	caaaagccagttatgaaaaaalttaattccatccagaattagataatgataagtc	963
Db	541	CAAAAGCCAGTAGTAGTGAATAAATTAATTCATACAGCAATTAATGATGATTAAGC	600
QY	964	gaagaataaacaagtttggattatattttaagaacctgtagtgaagtattatagataat	1023
Db	601	GAAAGAAAAACAAGTTTGGATATATTTTAACGAACCTGATGCCAAAAGTATTAATAT	660
QY	1024	ttatggttcgtattttagaatacctaagtttataagaagcaggttgaacctgtctct	1083
Db	661	TTATGTTCTGTTATTATTGATATCTAGGTTTATCAAGCAGAGTTGAACCTTCTCT	720
QY	1084	gagcaagccgcctgaatggttcgcacatgaagaagaacaagataacgcagtaactaat	1143
Db	721	GAGCAAGCCGCGCTCGAATGCTGCCATGAAGAGCAGCAGATTAACGAGTAGTAAT	780
QY	1144	aatgagttacagttatgctctataacaagaactgcgtcaagaagattatacaatgaat	1203
Db	781	AATGAGTTACAGTTAGTCTATATACAAACCTGCTCAAGCAAGATTAACAATTAAT	840
QY	1204	gaatgtttccggtgctgcagcagcaactttaa	1233
Db	841	GAAATGTTCCCGGTGTCACACCAATTTAA	870

RESULT 4

LOCUS 7141 bp DNA BCT 20-DEC-1995

DEFINITION: E. coli H+ ATPase, alpha, beta, gamma, delta and epsilon, and integral membrane proton channel of D and K C subunit genes, complete cds.

ACCESSION M25464

VERSION M25464.1 GI:146318

KEYWORDS H+ ATPase; integral membrane proton channel.

SOURCE Escherichia coli

ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE 1 (bases 1 to 7141)

AUTHORS Kanazawa, H. and Futai, M.

TITLE Structure and function of H+ ATPase

JOURNAL J. Biol. Chem. 267: 40267-45-64 (1992)

MEDLINE Ann. N. Y. Acad. Sci. 652: 45-64 (1992)

FEATURES

Source

Location/Qualifiers

1..7141

/organism="Escherichia coli"

/db_xref="taxon:562"

165..557

/note="14 Kd ORF; putative"

/codon_start=1

/transl_table=1

/product="unknown protein"

/protein_id="AA83868.1"

/db_xref="GI:1128958"

/translation="MKNVGSYLSVRNVAKKLLIVQLLVIASGLFSLKDPEFAVSA

ISGLAVFLPVLFIEMFAMRHOATPAKGRVAMFARGEAFKVLIVLVLAVALIK

AVFLPVLVIVLVIVQILAPVAVINNG"

566..1381

/codon_start=1

/transl_table=1

/product="integral membrane proton channel F0 subunit A"

/protein_id="AA83869.1"

/db_xref="GI:146319"

/translation="MASENNTPPOVIGHHLLNLTDLRFESINDPOMPAPAFWTIND

SMPSVVLGLFLVLRSAKKAISGPGKFORIAELVIGFVNGSYADMTGKSKLIA

PLAFTIVVFLNMLDLPLDILPYIAEHVGLPALRVVPSAVNLTSLMALGVFTL

ILFATIMKGIIGFTKELRCSTFNNHAFIPVNLILEGVSLSKRVSIGLRIFGMVAG

CDS
 1428..1667
 /codon_start=1
 /transl_table=11
 /product="Integral membrane proton channel F0 subunit C"
 /protein_id="AA083870.1"
 /db_xref="GI:146320"
 /translation="MENLMDLLYMAAAYMGLAIGAIGIGILGKFLGGAANQPD
 LIPLRTOFPIVGLVDAIPMIAVIGLIVMEAVA"
 1729..2199
 /codon_start=1
 /transl_table=11
 /product="Integral membrane proton channel F0 subunit B"
 /protein_id="AA083871.1"
 /db_xref="GI:146321"
 /translation="MNLNATILGOAIAFVLFVLCMKYVMPPLMAIEKREIADGL
 ASAERHKLDLAKASATDOLKKAQAVITIEQAKRRSOLIDEKAAEQERKIV
 AQAQATEIEKRRAREELKQVAIILVAQAEKIIERSVDEANSDIVDLVAVEL"
 2214..2747
 /codon_start=1
 /transl_table=11
 /product="H+ ATPase F0 delta subunit"
 /protein_id="AA083872.1"
 /db_xref="GI:146322"
 /translation="KSEFTVAPPYAKAAPDFAVEHOSVERMODMAFAEYTKRNM
 AELISGALPETAESFIAVCGEOLDENQNLIRVAENGRNLADPVLQFTHLAV
 SEATAEVDVISAALSEQOLAKISAMERLSRKVLNCKIDKSNVAGVIRAGDMVI
 DGSVGRLERLADVLQS"
 2760..4301
 /codon_start=1
 /transl_table=11
 /product="H+ ATPase F1 alpha subunit"
 /protein_id="AA083873.1"
 /db_xref="GI:146323"
 /translation="MOLNTEISELILKORIAFQNVYSEAHNEGTVISQGVIRIGL
 ACDMGEMISLPENRRAIALNLEPRDSVAVMGPYADLQMGVKGIRILEYVPG
 LIGRVNVLAPIDGKPLDHFSAEVALPAYERIOSDOYQYGRKRVDSMIRP
 RGRELIGDROTKTALADALIINORDSGIKOIVYALGKASTISNVKLEHGL
 ANIIVVAFAESAALOYLARFVALMGSEYFRGDEBALIYDLSKQAAVYQISL
 LRPPGEAPFPGVFLYHSLRLERAAVNAVEAFYKGVKGTSLALPIETEDG
 GVSATVPNTVISITDGOIFLETNLEAFGRPAVNPISYRVGGAOTIKMLSGG
 IRRALQVRELAFSOPASDLDDATRKQDHQKVELLKOKYAPMSVNOQSIYLEA
 AEGYIADVELSKISFEFALLAYVDRDHPAPMOEINQIGYINDEIGKLKILIDSK
 ATQSW"
 4352..5215
 /codon_start=1
 /transl_table=11
 /product="H+ ATPase F1 gamma subunit"
 /protein_id="AA083874.1"
 /db_xref="GI:146324"
 /translation="MAGAKETRSKIASVONTOKITKAMENVAASKMRKSDRMAASRP
 YAEKMRVYIHLAHNGLEYKHPYLEDDVYKRVGLVYSTDGCGGAINILFKLLAE
 MKMTDKVOCODLANLGSEKGVSPENSVGNVAOVGMDNPISLSEILGKVLQAY
 DEGLDKLYIVSNKFTIMTMSOVPITISOLLPLPSDDDDIKHMSDYLYEPDKLLLT
 LLKRVESQVYGGVENLASEQAAARMVAMKAADINGSLIKEQLYVINKRQASITOE
 LLEIVSGAAAV"
 5242..6624
 /codon_start=1
 /transl_table=11
 /product="H+ ATPase F1 beta subunit"
 /protein_id="AA083875.1"
 /db_xref="GI:146325"
 /translation="MATGKIVOVIGAVVDEFPDADVPVYDALEVONGNRLVLEVO
 OGLGGIVRTIAGMSDGLRGLDVKDLNHP IEVYVSKATIGRLINVLIGEVDKGET
 GEERMAIHRAPSYEELNSOELLEIGVITIDMCPAGGKVGILGGAGVGTVM
 MELIRNIAIHSIGTSVFAVGERTREGNDPHEMTDSNVIDKYSLYVGNMPEPNNL
 RVALTGLTMAEKFDEGRDVLLEVDNITRYTLAETFSALGRMPSVAGIOTPIAEEA
 GVLDERITSTETGSITVSQAVYPADNLTSPPTFAHDAATVLAHQSLISLITYA
 VDPIDTSRDLPLVGOEHDITARVOSIOTRQEKDLIAIIGMSEDLSEDLVVA
 RAKTORLSQPFVFAEFTGSPGKYVSLKDTINGFPGIMGEVDHLPQGFVYWGSI
 DEAVEKAKKL"
 6645..7064
 /codon_start=1

Query Match 41.2%; Score 812; DB 2; Length 7141;
 Best local Similarity 64.7%; Pred. No. 5,5e-149;
 Matches 1266; Conservative 0; Mismatches 660; Indels 48; Gaps 4;
 BASE COUNT 1679 a 1796 c 1966 g 1700 t
 ORIGIN 213 bp upstream of HindIII site.
 /transl_table=11
 /product="H+ ATPase F1 epsilon subunit"
 /protein_id="AA083876.1"
 /db_xref="GI:146326"
 /translation="MAMTYHLDVVAEQDMQESGLVEKIQVTSBSELGIYEPHAPLLT
 AIKPMIRIVKQHGHEEFYLSGILEVOPGWTVLADPTALINGOLDARAAEAKRKA
 EEHSSHGDDVVAQASAEIARAKIOLRIEILTKRAM"
 11 ttgctcagttcggcttcgattcctatgatgcacacgtaagaattatcacatgtcaa 70
 3975 ttctctcagtttcgacccgaccccttgacgatgacacacgtaagcagcttgaccacggtctacg 4034
 71 aaagtaactgaattatgaacaaacaaacaaactctccgtctatctgtagacaacaagca 130
 4035 AAGGTGACCGAAGTCTGGAACACAAACAGATATGCCCGATGTCGTCGGCAGCACTCT 4094
 131 ttatgttatttcagtagagtttggttacttagaagaatggaactagatcgatttgt 190
 4095 CTGTTTCTGTTCCGACGACGACGACGATGTTACTGGCGGATGTTGAATGTCGAAATGTGCC 4154
 191 tcatgtaactgaacttttagatgatacgcataactatgctgattttagctgtgag 250
 4155 AGCTTCGACGCCGCTCTGCTGCTTACGTCGACCGTATACGCTTCGTTGATGCAAGAG 4214
 251 ttaaccaactcgtgcaattcgaatgataaattaaagaatcaatgaatgatttgat 310
 4215 ATCAACCAAGCCGGTCTCAACGACGAAATCAAGCAAGCTGAAGAGCATCTCTGAT 4274
 311 agctcaagaacaaacagtcgtgta-----agtaact 346
 4275 TCCTTCAAGCAACCAATCCTGTTACGTCGCGCTTCTTGAAGCGAGCCGCAAG 4334
 347 ttaatgtagagacaaatgtagagttcctaagaagataagaacaaatcgtagttaa 406
 4335 GCATTGAGGAGAACCTGATGCGCCGCAAAAGAAATGATGAATGCAAGCGCTCC 4394
 407 aaagtaacaaanaaattactaaagcgaatggaatggtgtcgtcgcgaagaatcgtaaaa 466
 4395 AGAAGCGCAAAAGATCACTTAACGATGAGATGTCGCGCTTCCAAATGGGTAAAT 4454
 467 cgaagaacgcatgtctcttctcaacgaccttaacgaacaaatcgaatgtagcc 526
 4455 CGCAGGATCGATGGCGGCCAGCGCTCTTATGCAAAACATGCGCAAAATGATTGGTC 4514
 527 acgtttccaaagcgaatggttaacagatccaatttttagttagtcggaagtaaaa 586
 4515 ACCTTGCACACGGTAATCTGGAATATAAGCACCTTACTGTGAAGCGCAGCGTTAAAC 4574
 587 aagtggaatgattgtgtgtccacaagatcgtgtccttctgtgtggtcgaatgtaac 646
 4575 GCGTGGCTACCTGGTGGTGTGACCGACCGTGTGTTGGGGTGTGTAACATTAAC 4634
 647 tgttaaacctgtatlaaataaataaagaatggaagaagaagaatggtttccgttcaat 706
 4635 TGTTCAAAACCTGTCGGCGGAATGAAGACCTGGAACGCAAAAGCGTTCGAATGCGCAC 4694
 707 tgaatttaacgtgttctaactatcaacttttccaaatcttgggaattaaatttaa 766
 4695 TCGCAATGATCGGCTCGAAAGCGCTGCTTCTTCACTCCGTGGCGGCAATGTTGTTG 4754
 767 cccaagattcaagatgtgtgatactccctcgtgtgaacagtaatttggtaagcaatt 826
 4755 CCCAGTCAACCGGATGGGGGATTAACCTTCCTGCTCCGAATGATCGCTCGGATTAAG 4814
 827 ctatgatgattgctataaaaagggaagtagatgtgtgtatttagttaacaat 886

D	4815	TGATGTTGACAGGCTTACGACGAAGGCCGTCGTGACACAGCTTTTACATTGTGACGAACAAAT	4874
O	887	ttataaacagatggtcgcgaaaagccagta ttgaaaaataattccattaccagaattag	946
D	4875	TTATTAAACCAATGTCTCAGGTTCCGACCATCGACGAGCTGCTGCCGTTACCGGATCAG	4934
O	947	ataatgtgaattagcgcgaagaacaagtttggattatattacgaacctgatcga	1006
D	4935	ATGATGATGATCT-----GAAACATTAATCCGTGGATTACTGTCAGCAACCCGATCCGA	4988
O	1007	aagattattagataattttagtctgtatttagaattcctaagtttataagaagag	1066
D	4989	AGCGGTTGCTGGATACCTCTCTCGTCGTTATGTGCAACTTAGTTATACAGGCGGG	5048
O	1067	ttgaaaccttgctctgtgaacagccgtctcgaattgtctgcca ttgaagcaacaga ta	1126
D	5049	TTGAAACCTGGCCAGCAGCAGCGCCCGATAGTGTCGATGAAGCCGAGCCGACGA	5108
O	1127	acgcagtaacttaataatgatgattacgttagtctataaagaacctcgtaagca ta	1186
D	5109	ATGCGGCACCGCTGATTTAAAGACTCTCAGTTGGTATACAAACAACTGTCAGCCACGA	5188
O	1187	ttacaatgaataaatagaattgttcgcgggycgagcaatttaaca-----a	1237
D	5169	TTACTCTAGGAACCTACCGAATGCTCTCGGGGGCGCCCGGTTTAAACAGGATATTTTCG	5228
O	1238	taagaga tgcgtaatgycacatctgaaaaattgacaacatcactcgtgcgtattagct	1297
D	5229	TAGAGGATTTTAAGATGCTACTGTGGAAGATTGTCCACGTAATCGGCCCGCTAGTTGACGT	5288
O	1298	tgaattccacaagatgcagtaaccaaaagtatatgtgcct-----aaatgtga	1348
D	5289	CGAATTCCTCCAGAGATGCCCCCTACCGCGCTGTACAGTCTCTTGAGGTGCAAAATGTAA	5348
O	1349	aacaggttagtacttgaagttaacaaacaattagtggtgtgtagtgcgtatgc	1408
D	5349	TGACCGTCTGTGTGTGGAAATTGACGACAGCTCGCGCGGGGTATCTGACGACCATATGC	5408
O	1409	aatgga tca tctgatagtatlaaagcggttttaag cgtacaacaa taacgaataccaat	1468
D	5409	AATGGTTCCTCCGAGCGTGTGCTGCGGCTGTGATGTAAAGACCTCGAACACCCAT	5468
O	1469	tctcgttccagtggaacgaaaaacttggctcgatcabyaacgta tttggttgaaccaat	1528
D	5469	TGAAGTCCCGGTAGGTAAACCGACTCTGGGCGGTATCATGACGATCTGGGTGAACCGT	5528
O	1529	cgaatgacgaagtgaataatcggtgcgaagaagaatgtgctatataccgtgcgcaaccaag	1588
D	5529	CGAATTAAGGGCGAATCCGTAAATAAAGATGGTGGGATTCACCCGACGACCTTC	5588
O	1589	ttatgaagaacaatcacaagtaacttttagaaaaggaattaaagtatcgaact	1648
D	5589	CTAGGAAGACCTGTCAAACTCTCAGGAACCTGCTGGAAACCGGTATCAAGTTATGACCT	5648
O	1649	agtttgcgttctgcgaaggggttaagttatcgttgcgttgcgggtgcgttgcgtta	1708
D	5649	GATGTGCCGTCGATAAGGGCGGTAAAGTTGGTGTGTTCCGTGGTGGGTGATAGTTAA	5708
O	1709	aaccgtcaataatgattgaattaatccgtacaatcgcgaattgaacactgaattactctgt	1768
D	5709	AACCGTAAACATATGAGAGCTACTTCGTATACATCCGATCCGAGCACTCCGGTTACTCGT	5768
O	1769	ctttgcggggtagtgtagcggtacgtgtaaggttagaacttctatactgtatgaaaga	1828
D	5769	GTTTGGCGGGTAGGTGAAGTACTGTGAAGGTAAACGCTTACACGAAATGACCGA	5828
O	1829	ctctaaagtaattagataaagtgctctcttattatgcaatgaaacgagccacaagta	1888
D	5829	CTCCAAAGTTATGACAAAGTATCCCTGTGTTATGCCCCAGATGAACGAGCCCGGGA	5888
O	1889	ccgtttacgtgtgcatlaacaggttlaactatgycggaanaatccgtgtatgaagtcg	1948
D	5889	CCGTCGCGGTTGCTCTGTGACCGGCTTGACATAGGCTGAATAATCCGTGACGAAGGTGC	5948

QY	1949	tgatgtcttattcttcgttgataa	1972
Db	5949	TGACGTTCTGCTGTTGTTGACAA	5972

RESULT 5

LOCUS	11414 bp	DNA	BC ¹	01-DEC-2000
DEFINITION	Escherichia coli K12 MG1655 section 340 of 400 of the complete genome.			

VERSION

SOURCE	ORGANISM
Escherichia coli K12.	Escherichia coli K12

REFERENCE

REFERENCE
AUTHORS

1 (bases 1 to 11414)
Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V.,
Riley, M., Collado-Vides, J., Glasner, J.D., Koeck, C.K., Mayhew, G.F.,
Gregor, J., Davis, N.M., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J.,
Mau, B. and Shao, Y.

JOURNAL

MEDLINE	97426617
PUBMED	9278503
REFERENCE	2 (bases 1 to 11414)

TITLE

JOURNAL
Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecollegenetics.wisc.edu Phone: 608-262-2534 Fax:
608-262-7450

REFERENCE

AUTHORS Bialchev, F.R.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.

REFERENCE	4 (bases 1 to 11414)
AUTHORS	Plunkett, G. III.
TITLE	Direct Submission

1
2
3
4
5
6

COMMENT This sequence was determined by the E. coli genome project at the University of Wisconsin-Madison (Frederick R. Blattner, director). Supported by NIH grants HG00301 and HG01428 (from the Human Genome

Project and NCBI). The entire sequence was independently determined from E. coli K12 strain MG1655. Predicted open reading frames were determined using Genemark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332 (e-mail: mark@amber.gatech.edu). Open reading frames that have been correlated with genetic loci are being annotated with CG Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berryn. A public version of the database is accessible (<http://cgsc.biology.yale.edu>). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (<http://www.genetics.wisc.edu>). *** The E. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.

Dh 7035 AAGTGAACGACGACGCTGTAACGAAACAGTATGCCGCGATGCTCCGTTGCGCAGCAGTCT 6976
Qy 131 ttatgtatatttcgagtagagtttggttacttagaagaagttagtacttgatcgtattggt 190
Dh 6975 CTGGTTCGTTCTCGCAGCAGAACGTTGTTACTGCGGATGTTGAACCTGTGCAAAATGTGCG 6916
Qy 191 tcaattgatacgaacactttagatgatacgaatacattatcgtattttagcgttag 250
Dh 6915 AGCTTCGAAAGCGCTCTGCTGCTTACGTCGACCGTGATACAGCTCCGTTGATGCAAGAG 6856
Qy 251 ttaaccaactctgagcaatlacaaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 310
Dh 6855 ATCAACCGACCGGCTGCTGCTACCAACGAAATCGAAGGCAACCTGAAGGCACTCTCGAT 6796
Qy 311 agcttcaagaagaacagtgctggtgta-----agttacact 346
Dh 6795 TCCTTCGAAGCAACCAATCTGCTGATGCTGCGGCTTGGCTTGAAGGCGCGCAAG 6736
Qy 347 ttaaatgagagacaaatctgagcagtgctaaagagataagaacaaatcgcagtgtaa 406
Dh 6735 GCATTCGAGGAGAGAGCTCATMGCGCGCGCAAAAGATACGTATGATGCAACGCTGCC 6676
Qy 407 aaagtacaaataaactctaaagcagatggaatggtgctgctcctgaataatgctgtaaa 466
Dh 6675 AGAACCCGCAAAAGATCACTAAAGCGATGAGATGTCGCCCTTCCAAATGCGTAAT 6616
Qy 467 cgaagaagcagatgcttcttcaagccttattcagaacaaatgaatgaatgaatgctgctg 526
Dh 6615 CGCAGGATCGCAGTGGCGCGCAGCCGCTCTTATGCAAAACCTGCGCAAAAGATGTCG 6556
Qy 527 acgttccaagaacagatctggttcaagcattcattttagtgatgagcgaagttaaa 586
Dh 6555 ACCTTCGACAGGATGATGGAATTAAGCAACCTTACCTGGAACCGCGACGTTAAAC 6496
Qy 587 aatgagcagatgctgctgctcagacagatcgtgctgctgctgctgctgctgctgctgctg 646
Dh 6495 GCGTGGCTTACCTGCTGCTGCTGACCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6436
Qy 647 ttttaaaactgatttaataaagaagaatggaagaagaagaagaagaagaagaagaagaagaaga 706
Dh 6435 TGTTCAAAACCTGCTGGCGGAAATGAAGACCTGAGACGCAAAAGCGCTCAATGCGACC 6376
Qy 707 tgaatttcaatcgttcttcaatctatcaacttlltcaacttcttggaatttcaatcttca 766
Dh 6375 TCGCAATGATCGGCTGCAAGAGCGCTGCTTCACTCACTCGTGGCGGCAATGTTGTTG 6316
Qy 767 cccaagatcagatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 826
Dh 6315 CCCAGGTCACCGGATGCGGATTAACCTTCCCTGCTCGAATGATGCTGCTGCTGCTGCTGCTG 6256
Qy 827 ctatgattgattctataaagaaggaagaatgattgattgatttatttataaacaat 886
Dh 6255 TGAATGTCAGGCTTACACAGAGCGCTGAGACAGCTTTTACATGTCAGCAACAAT 6196
Qy 887 ttaataacagatgctgcaaaagcagatggaataaataatcattcattcagatgattag 946
Dh 6195 TTAATTAACACCATGCTCAGTTCCGACCATCAGCACCTGCTGCTGCTGCTGCTGCTGCTG 6136
Qy 947 ataagtatgaattgagcgaagaagaagaatggaattggaattataattagaacctgattcga 1006
Dh 6135 ATGATGATGATCT-----GAAACATTAATCCTGAGATTACTGTTGAAACCCGATCCGA 6082
Qy 1007 aagattattttagataatttatttatttatttatttatttatttatttatttatttatttattt 1066
Dh 6081 AGCGGTCCTGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6022
Qy 1067 ttgaaacctgcttctgcaagcagcctgcaatgctgcaatgctgcaatgctgcaatgctgcaatg 1126
Dh 6021 TTGAAAACCTGCGCAGCAGAGCGCGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5962
Qy 1127 acgcaagtaacttaataatgagttacagttagttctataacaaagctgctgcaagcaagta 1186
Dh 5961 ATGCGCGCAGCTGATTAAGAGCTGCAAGTTGATATACAAAGCTGCTGAGCGCACCA 5902

Qy 1187 ttacaaatgaatlaataatgaattgttcgcggtgcaagcagcaatttaacaa-----a 1237
Dh 5901 TTACTCAGGACTCACCGAGATGCTCGGGGGCGCGCGCTTTAAACGATTAATTCG 5842
Qy 1238 tagagatcgttaatggaacactggaataatgtacaaatcagctgctgattttagcgt 1297
Dh 5841 TAGAGATTTAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5782
Qy 1298 tgaattcccaagaatgcaatcccaagaatataatgattgctt-----aatgttga 1348
Dh 5781 CGAATTCCTCAGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5722
Qy 1349 aacagtttagtacttgaatgattcaacaacaatgaatggtggtggtggtggtggtggtg 1408
Dh 5721 TGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5662
Qy 1409 aatgagatcatctgattgattaaagcgtttaaagcgtttaaagcgtttaaagcgtttaaag 1468
Dh 5661 AATGGGTTCTCCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5602
Qy 1469 ttcgttccagtgaggaaacaaatggttgcgtatcatgaaacgtatggtgaaacaaat 1528
Dh 5601 TGAATCCCGGATGAAAGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5542
Qy 1529 cgaatgagcagtgaaatcgcgtgcaagaagaatggttctatcaccgtgctgcaacaaag 1588
Dh 5541 CGACATGAAAGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5482
Qy 1589 ttaagaagaacatctaaacagatcagcattttagaagcaggaatlaaagttatgact 1648
Dh 5481 CTACGAAGCTGCTAAACTCTCAGGAATGCTGGAACCGGTATCAAAAGTTATGACCT 5422
Qy 1649 agttgcgcttgcgaaggggaaagaaagaaagaaagaaagaaagaaagaaagaaagaaag 1708
Dh 5421 GATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5362
Qy 1709 aacgtaataatgagaaatlaatcgttaacatcgaatgcaatgcaatgcaatgcaatgcaat 1768
Dh 5361 AACCGTAACATGATGAGCTCTATTCGTAACATGCGATGCGACCTCGGTTACTGCT 5302
Qy 1769 ctctgagggggaagtgagcagcgtgaaagtgaaagtgaaagtgaaagtgaaagtgaaagtg 1828
Dh 5301 GTTTCGGGCGGTAGGTGAACGTAACGTAACGTAACGTAACGTAACGTAACGTAACGTAAC 5242
Qy 1829 ctctaaagcattagaatgctgctgctgctgctgctgctgctgctgctgctgctgctg 1888
Dh 5241 CTCAACCTTATCGCAAAAGTATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5182
Qy 1889 cgtttagctgctgagcattaaagcgttactatgacggaataatcgtttagaagctg 1948
Dh 5181 CCGTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5122
Qy 1949 tgaatgcttattcttctgctgataa 1972
Dh 5121 TGACGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5098

RESULT 6
ECUNC 14526 bp DNA BCT 29-APR-1998
LOCUS
DEFINITION E. coli origin of replication oric and genes gid, unc, EcoRI-F1 and
gms.
ACCESSION X01631
VERSION X01631.1 GI:43256
KEYWORDS inverted repeat; origin of replication; transposable element.
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
REFERENCE 1 (bases 1 to 14526)
AUTHORS Walker,J.E., Gay,N.J., Saraste,M. and Eberle,A.N.
TITLE DNA sequence around the Escherichia coli unc operon. Completion of

the sequence of a 17 kilobase segment containing asnA, orfC, unc,
glns and phos
Biochem. J. 224 (3), 799-815 (1984)

JOURNAL
MEDLINE
REFERENCE

2 (bases 1 to 14526)

AUTHORS

Messer, W., Hartmann-Kuehlein, H., Langer, U., Mahlow, E., Roth, A.,

TITLE

The complex for replication initiation of *Escherichia coli*

JOURNAL

Chromosoma 102, 1-6 (1992)

FEATURES

location/Qualifiers

SOURCE

1. 14526

CDS

/organism="Escherichia coli"

CDS

/db_xref="taxon:562"

CDS

/note="16k protein"

CDS

/codon_start=2

CDS

/transl_table=1

CDS

/protein_id="CAA25772.1"

CDS

/db_xref="GI:809693"

CDS

/db_xref="SWISS-PROT:P03817"

CDS

/translation="GSVNLK"

CDS

/note="min. region of orfC"

CDS

/note="pot. promoter -35 region"

CDS

/note="pot. promoter -10 region"

CDS

/note="pot. Shine Dalgarno region"

CDS

/note="gida protein"

CDS

/codon_start=1

CDS

/transl_table=1

CDS

/protein_id="CAA25773.1"

CDS

/db_xref="GI:43257"

CDS

/translation="MYDPEDVYIIIGGSHAGTEAAMAARMGQTLITLHNIDPLGQ

CDS

LYROAVFRLLENQNLIMFQDAVEDLIVENDRYGAVTQGLKFRKAVAVLVGVFLD

CDS

KGTHGLDNYSGGAGDPSISPLSRRLPLTVGLKGTGPPRIDARTIDFVSVAQO

CDS

HNDNRPVSPMGNASOHPOQVPCYITTHNEKTHDVRISNDRSPMYAGVIGVPRY

CDS

CSYIEKVRKFRADRNHOJFLPEBGTSNEIYPNGISLSPDYMOQIYRSQGMENA

CDS

KIVRQYALEYDFDPRDLPTLESKFITGLKEPRYRFTSRAEYRLMRDNADRLITE

CDS

IGREGLYDDERMARFNEKLENIERERQRLKFTWVPSEAAAEVNAHLTAPLSRAS

CDS

GBDLRPENTYERKLTITLTPFAVALTDEQAAEQVEIQYKEGYIARQODIEKQIREN

CDS

TLLPRLDXYRQYSGLSNEVIATKNDHKPASIGASRISGVTTPAAILLWTLKQGL

CDS

RRSA"

CDS

2305. .2330

CDS

/note="pot. hairpin structure"

CDS

2346. .2349

CDS

/note="pot. Shine Dalgarno region"

CDS

2357. .2980

CDS

/note="gida protein"

CDS

/codon_start=1

CDS

/transl_table=1

CDS

/protein_id="CAA25774.1"

CDS

/db_xref="GI:581247"

CDS

/translation="MUNKSLIKDAGISLTDHOKNOLAVYMLLHKMNKAYNLTSVR

CDS

DPNEMLVRLHLDSTIVAPYIQGERFLDVGTGPGPLPGIPSLISVREAHFTLDSIGRV

CDS

RFLROYQHELKLEIEPQSVSEPEPDSVISAFAASLNDMSWCHHLGEBOGR

CDS

FVALKGMPEDEIALPEEYQVESVVKLQVAPALDGERHLVIAKANKI"

/protein_id="CAA25775.1"

/db_xref="GI:581248"

/translation="SWISS-PROT:P03808"

/translation="MKNVMSYLSVSRVARKLLVOLLVYIASGLFSLDPDWGVSVA

ISGLAVFLPVLVLFMIFAMRHQARPAKRAMFFAFGEAKFVLAAMLVLVALAVLK

AVFLPLVTVVTVLVOILAPAVINKG"

3979. .3981

/note="pot. Shine Dalgarno region"

3986. .4801

/note="uncB protein"

/codon_start=1

/transl_table=1

/protein_id="CAA25776.1"

/db_xref="GI:43260"

/db_xref="SWISS-PROT:P00855"

/translation="MASENMTPODYIGHHLNNLODLRTFSLVDPQNPATFWTINID

SMFSSVLGLFLVLFPSYAKKATSGVGRKOTFELIVGPNVSGKMDYKSKSLIA

PLALITFWVFLMIMLMDLLPIDLPIYIEHVLGIPALRVPSADYNWTLISALGVFL

ILFYSIKMKGIGGFTKELIQPFNMWAFIPNLILEGVSLSLKPSVSLGRFLGMYAG

ELIFLIGLPLMWSOWILNVPMAIFHILITLQAFIFMVLLIYLSMASEH"

4838. .4841

/note="pot. Shine Dalgarno region"

4848. .5087

/note="uncC protein"

/codon_start=1

/transl_table=1

/protein_id="CAA25777.1"

/db_xref="GI:43261"

/db_xref="SWISS-PROT:P00844"

/translation="MENLMDLITMAAVVMKGLAIGAIGIGIGKFLBEGANRPD

LPLRLTQFFIVMGLVDAIPMTAVGLGYMFAVA"

5146. .5149

/note="pot. Shine Dalgarno region"

5149. .5619

/note="uncF protein"

/codon_start=1

/transl_table=1

/protein_id="CAA25778.1"

/db_xref="GI:581249"

/db_xref="SWISS-PROT:P00859"

/translation="MNLNTITIGQATFVLFQCKKQYVPLMAAIEKREKELADGL

ASAEARHKLDLAKASATDQLKAKAEQVITIEQANKRSQILDEAKAABDEERKIV

AOQAQIEIEKRRARELRKQVAILVAQAETIERSVDEAANSIDVLKVAL"

5619. .5622

/note="pot. Shine Dalgarno region"

5625. .5628

/note="pot. Shine Dalgarno region"

5634. .6167

/note="uncB protein"

/codon_start=1

/transl_table=1

/protein_id="CAA25779.1"

/db_xref="GI:43263"

/db_xref="SWISS-PROT:P00831"

/translation="MSEFIVARPYAKAFDFEAVERHOSVERMODMATAFAEYTKNEOM

AELLSCALPETALESFIAVCGOLDENGONLIRVAENGRNALADPVLEQFIHLAV

SEATAVDIVISAALSEQOLAKISAMERKLSRKVLNCKIDKSVAGVIRAGDMYI

DGVRGRTERLADVLOS"

6175. .6178

/note="pot. Shine Dalgarno region"

6180. .7721

/note="uncA protein"

/codon_start=1

/transl_table=1

/protein_id="CAA25780.1"

/db_xref="GI:43264"

/db_xref="SWISS-PROT:P00822"

/translation="MQLNSTEISELTKQRIAQPNVYSEAHNECTIVSVDGYRIHGL

ADCMOGEMISLPGNRVAILNLBERDSVAVMGPPYADLAGMKVYKCTRIIEVPVGG

ILGRVYVTLGAPIDGGLDHPDSVAEALAVGIEROSVDOPVOTGKAVDSMIPIG

RGQRELITGRQGTKRLAIDAIINQDSIKTIYVAGIKASITISNVKLEHGL

ANTIVVATSESAALOYLARMPVALMGEFFRDRGDEALITIDYDSQAAYROIISL

LRRPGREAPGDFVYLDHSRLERAAVNAVEAFYKGGKGTGSLTALPIETQVA

GDVSAFVPTNVISITDQIFLETNLFNAGIRPAPVNPGISVRGAAQTKIMKLKSG
 IRRYLAQYRELAASFOSAPSDLDATRKOLDKROLDGVTTELKOKOAPSVAAQOSIYVFA
 AERGLADVELSKISGSFEALALAYVDHBMIMQEIINGTGSYNDIEGKIKGIIDSRK
 ATOSM"
 RBS 7726..7753
 /note="pot. hairpin structure"
 7761..7764
 /note="pot. Shine Dalgarno region"
 7772..8635
 /note="uncd protein"
 /codon_start=1
 /transl_table=1
 /protein_id="CA25781.1"
 /db_xref="GI:43265"
 /db_xref="SWISS-PROT:P00837"
 /translation="MAGAKETRSKIASVQNTQTKTKMENVYASKMRKSODRMAASRP
 YAEIWRKIVIGLHAGNLEKHPYLEDKRVKGLVYSTDGLGGININLFFKLAE
 MKWTMDGVOCDLAMIGSKVSPNSGVGNVAVTGMDNPSLSEILGIPKVALDQY
 DEGRDLXLYISNKFNTMSQVPTISOLLPLPADDDDKHKSMDYLYEPDKALDPT
 LLRRVSOYQGVENVLASBQAAKRMVAKMATDNGSLIKELQLVYNNKARQASITQE
 LLETVSGAAAV"
 RBS 8652..8655
 /note="pot. Shine Dalgarno region"
 8662..10044
 /note="uncd protein"
 /codon_start=1
 /transl_table=1
 /protein_id="CA25782.1"
 /db_xref="GI:899257"
 /db_xref="SWISS-PROT:P00824"
 /translation="MATGIQVIVGAVYDVEFPDQAVPRVYDALEVONGNERLYEVO
 OULGGIVRTIAMSGSDGLRGLDKLEHIEVPKATKIRIMNYLGEFVPMKKEI
 GEERPMRIHRAAPSEELNSQELLEIKYIDLMCPKAPKAGKGLGSGAVGTVNM
 MELRNATIHSGISVFAGVGEKREBNDHEHTDSVNDKSLVVGQMPGNRL
 RVALTGLTMAEKFEDEGRDVLLEVDNITRYTLAEVSALLRMP-SANGVOPTLAEM
 GVLOERTSTRTGSIYSQVAVPADLTPSPATFAHLDAFVLSKQISLGIYPA
 VDPIDSTRDPLIVGOEHDTARGVOSILOROEIKDIIAILGMDLSEEDKLVA
 RARKIORFLSGPFFVAEFTGSPKYSIMKTINGFIMGEDHLPQDAFYVWGSI
 EEAVERAKKT"
 RBS 10057..10059
 /note="pot. Shine Dalgarno region"
 41.2%; Score 812; DB 2; Length 14526;
 Best Local Similarity 64.7%; Pred. No. 5.3e-149;
 Matches 1296; Conservative 0; Mismatches 660; Indels 48; Gaps 4;

407 aagttacacaaaataactaaagcgatggaaatggttcgtccctcgaaatgcgtaa 466
 7815 AGACACGCAAAAGATCTACTAAAGCGATGGAGATGGTGGCCCTTCCAAATGCTTAAT 7874
 467 cgcagaacgcgatgtctcttcacgccttaccagaacaatacgaatcgaatgtagcc 526
 7875 CGCAGATCGCATGGCGGCGCACCGTCTCTTATGACAGAAACATGCGCAAGATGATGGTC 7934
 527 acgttccaagaacgatgtgttacaagcatccattttagtgaatcggaatgtaaaa 586
 7935 ACCTTGCACACGATATCGAATTAAGCACCTTACTGTGAACCGCGCATTAAC 7994
 587 aagttggatgattgtgttcacacagatcggtgttggtgtgtttaaagcgtgaact 646
 7995 GCGTGGCTACCTGGTGTGTGACCGACCGGTTGTGGCGGTGTGAACATTAAC 8054
 647 tgtttaaactgtatataatgaatgaagaatggaagaagaatgtttccgttcaat 706
 8055 TGTTCAAAACCTGCTGGCGGAATGAAAGACTGGACCGACAAAGCGGTCAATGCGACC 8114
 707 tgaatttaactcgttctaataatcacaacttttccaacttttgggaattaaatttaa 766
 8115 TCGCATATGATCGCTCGAAAGCGGTGTCTTCACATCTCGTGGCGGCAATGTTGTG 8174
 767 cccaagattcagatattgtatcactccctcgtgtgagcaatgattgttcaatcaat 826
 8175 CCCAGGTACCGCGCATGGGGGATTAACCTTCCCTCCGAACTGATCGGTCCGTAANAAG 8234
 827 ctatgatgatcctataaaaaagggaagatgattgtgttattttagtttatacaaat 886
 8235 TGATTTGACAGCCATACGACGAAAGCCGCTGGAACAAGTTTACATTTGACACAAAT 8294
 887 ttattaacagatgcgcaaaaagccagatattggaataatlaattcattaccagaattag 946
 8295 TTATTAACACCATGCTCTAGGTTCCGACATCAGCACCTGCTGCCGTTACCGCATAG 8354
 947 ataagatgaattagcgaagaagaagaagtttggatattataatgaactgtatgga 1006
 8355 ATGATGATGATCT-----GAAACATTAATCCTGGATTAACCTGTACGAAACCGATCGGA 8408
 1007 aagttatagaataattatgttcgttatttagaatcctcagttatacaagaagcag 1066
 8409 AGCGTTGCTGATACCTGCTGCGCTGTTATGTGATGATGATGATGATGATGATGATGAT 8468
 1067 ttgaaacctgtcttcgaagaacgcgtcgaaatgtgcgaatgaagaagcagaagata 1126
 8469 TTGAATAACCTGCGCAGCAGCAGCGCCGCGGTATGTCGATGAAGCCGCGACGACGA 8528
 1127 acgcagtaacttaattagatgaatgtagttagtctataaagaactgtgaagaagta 1186
 8529 ATGGCGGCGCTGATTAAGAGCTGCAAGTTGGTATACAAACAACTGTCAGGCGACGA 8588
 1187 ttacaatgaatlaaataatgttgcggtgcgagcagaatlaaaca-----a 1237
 8589 TTACTCAGGAACCTACCGAGATCGTCTCGGGGCGCGCGGTTAAACAGTATTTTCG 8648
 1238 tagagatcgtgaatgtgcaactggaataatgtacaaatcactcgtgttatgtgact 1297
 8649 TAGAGAGATTTAAGATGCTGACGGAAGATGTCAGGATTAACGCGCGCGATGAGAGT 8708
 1298 tgaattccacaagaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 1348
 8709 GCAATTTCCCTCAGATGCGGTACCGCGCGGTGATGATGATGATGATGATGATGATGAT 8768
 1349 aacagtttagtacttgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 1408
 8769 TGAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8828
 1409 aatggatcattcgtatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 1468
 8829 AATGGATTCCTCCGAGGCTGCTGCGGTGATGATGATGATGATGATGATGATGATGAT 8888

QY	1469	ttctgttcacgttgggaacggaacacattgggtcgtatcatcagaagatttgggttaacaaat	1528
Db	8889	TGAAGTCCCGGTGTGGTAAACCGACTCTCGGGCCGTATCATTAACGTATGGGTGAACCGGT	8948
QY	1529	cgatgtgcaaggtggaatcgtgtgcagaagaattgtgttcattcaaccgttgcgcaccaaag	1588
Db	8949	CGACATGAAGGCGAGATCCGGTGAAGAAGACGTTGGGCATTCAACGGCAGACACTTC	9008
QY	1589	ttaatgaagaacaactaacagactgaactttagaacaagggaattaaagtatcagactt	1648
Db	9009	CTAGAGAGACTGTCAAACTCTCAGAACTGTGGAACCGGTATCAAGTTATCGACCT	9068
QY	1649	agttgtcgttttgcgaagaagggttaagtaagtatttcttgggtgtgcgggtgcgttaa	1708
Db	9069	GATGTCTCCCTTCCTTAAGGCGCGTAAAGTGGTCTGTGGTGTGTCCGGGTGAGGTAA	9128
QY	1709	aaccgttaatatgtatgtgaattaatccgttaacatcgcaatcgaactgaacacagttactcgt	1768
Db	9129	AACCGTAAACATGATGTGAGCTCATTTCTGTAACTGCGCATCGACACATCCGGTTACTCTGT	9188
QY	1769	ctttcggggggttaagttagtcgcgtgtgaaggttaacgacttctatcatagatgaaga	1828
Db	9189	GTTTGGCGGGGTAGTGTGACGTACTCGAGAGGTAAAGACTTTCACACAGAAAGACCGA	9248
QY	1829	ctctaacgtattagataaagtgtctcttgtttatggttcgaatgaagcagcagcaaggttaa	1888
Db	9249	CTCCAAAGGTTATCCACAAAGATCCCTGGTGTATGGCCAGATGAACGAGCCCGGGGAAA	9308
QY	1889	ccggttcggttgcgcattaaaggcttaacactatgacggaataatccgttgaagtcg	1948
Db	9309	CCGTCGTGGCGCTTCTCTGACGCGGTCTGACCATGCGTGAAGAAATTCCTGTACGAAGGTGCG	9368
QY	1949	tgaatgcttattcttcgttgaata	1972
Db	9369	TGACGTTCTGCTGTTCGTTCGTGACAA	9392
RESULT/7	ECOWM82/c		
LOCUS	ECOWM82	136254 bp	DNA
DEFINITION	E. coli; the region from 81.5 to 84.5 minutes.		
ACCESSION	L10328		
VERSION	L10328.1	GI:290484	
KEYWORDS			
SOURCE			
ORGANISM	Escherichia coli K12 strain MG1655; lambda clones EC14-52, EC17-187, EC27-890, EC21-99, EC22-175, EC27-236, EC17-30, EC19-71, EC15-15, EC17-159; subclones in M13mp19 or Janus. Escherichia coli Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.		
REFERENCE	1 (bases 1 to 136254)		
AUTHORS	Burand,V., Plunkett,G. III., Daniels,D.L. and Blattner,F.R.		
TITLE	DNA sequence and analysis of 136 kilobases of the Escherichia coli genome: organizational symmetry around the origin of replication		
JOURNAL	Genomics 16 (3), 551-561 (1993)		
MEDLINE	93315143		
COMMENT	This sequence was determined as part of the E. coli Genome Project (Frederick R. Blattner, director) at the University of Wisconsin-Madison. Supported by award HG00301 from the NIH Human Genome Project. A preliminary report was presented at the Hilton Head meeting Genome Sequencing and Analysis IV, September 26 - 30, 1992. The entire sequence was independently determined from E coli MG1655; overlaps and conflicts with other sequence determinations are annotated. The end of this entry overlaps the start of the entry ECOWM85 (M87049) by the six bases of an EcoRI site. Data kindly submitted in computer readable form by: Guy Plunkett III		
	Laboratory of Genetics		
	University of Wisconsin		
	445 Henry Mall		
	Madison WI 53706		
	USA		

Phone: 608-262-2534
Email: ecoligenetics.wisc.edu
Fax: 608-263-7459
[1] Authors request hold until publication
location/Qualifiers
1. 136254
source
"Escherichia coli"

misc_feature complement(<1. .669)
/note="X06036; ECOFPG(1. .>669)"
1. .11607
/note="EC14-52 (Ecofa clone)"
complement(<1. .544)
/gene="fpg"
/note="alternate gene name mutH"
/codon_start=1
/transl_table=11
/product="formamidopyrimidine-DNA glycosylase"
/protein_id="AA61988.1"
/db_xref="GI:643609"
/translation="MPLELPEGVTSRRGIEPHLVGATLLAAVVBNGRLRMPVEEIRL
SDQPLASVORAKYLLELPEGWIIHLGSGSLRIPELPEKDHVDVLSNKKV
LRYDPRRGFWMTQKEGHNVTLHGPPLSDPFNGEYLHQCAKKKTAIKPMLAD
NKLIVGNTIYASSLPRAGI"
complement(1. .544)
/gene="fpg"
complement(526. .1290)
/note="J01677; ECORPMBG(1. .764)"
complement(557. .586)
/note="promoter-like sequence; promoter matrix score of
44"
misc_difference 596. .601
/note="AAAAA in X06036; AAAAA in J01677 and here"
misc_difference 611
/note="G in J01677; A in X06036 and here"
misc_difference 615
/note="G in J01677; A in X06036 and here"
misc_difference 618. .623
/note="TTTTTTT in X06036; TTTTTT in J01677 and here"
complement(642. .809)
/gene="rpmg (CG site No. 236)"
complement(642. .809)
/gene="rpmg (CG site No. 236)"
/codon_start=1
/transl_table=11
/product="50S ribosomal subunit protein L33"
/protein_id="AAA61989.1"
/db_xref="GI:290486"
/translation="NAKGIREKIKLVSSAGCHFTTTKKNRTPKELKKPDPVVR
QHYVKEAKIK"
complement(830. .1066)
/gene="rpm (CG site No. 240)"
complement(830. .1066)
/gene="rpm (CG site No. 240)"
/codon_start=1
/transl_table=11
/product="50S ribosomal subunit protein L28"
/protein_id="AAA61990.1"
/db_xref="GI:290487"
/translation="MSRCOVGVGRRPYVGNNSHALNTKRRFLPNLHSHPVWESEK
REYTLVYSAKMRVYDKRIDIVLAEIARSGKY"
misc_difference 902
/note="G in J01677; A here"
misc_difference 1128
/note="TC in J01677; C here"
misc_difference 1153. .1155
/note="T in J01677; TAT here"
complement(1168. .1819)
/note="X63366; ECORADC(1. .651)"
complement(1212. .1240)
/note="promoter-like sequence; promoter matrix score of
77"
gene complement(1283. .1957)

ELGAAHVHLYVHGSGDILKQALKIDONI.MWVLAQDLYGNGHMQQAAPPEADDEDTLIL
 YEDVAHLYISERTIQRLRDAPKPOSGITELLVYKLDLDPGCRITRENGKYTGIEKHDDTLI
 EDQRQLOEINTGLILANGADMKWIAKLITNNAGGEYITTDITLALVDEGRITAVHQ
 RLSEGEVNNRLQLSRLERVYQSEAEKLLAGVALRDPARFDLRTGLTLHRDVEIDTDP
 NVIIEGNTVLGHRRKIGTGVYKNSVIDDCEISPYTVEDBANLAACTIGSPFARLRLP
 GAELLEGHAGVNGEYMKRKARLGKSGKASHLYLIDDAELIGDNVNIAGTITCYXGANKK
 FRTIIGDEYFVSQDTQLVAPLYTGKATILAACTYTRNVGENALALISNVPQTOEGRRR
 RPVKK"

gene
 complement(1834. .2253)
 /gene="atpc"
 /note="55229"
 cds
 complement(1834. .2253)
 /gene="atpc"
 /function="enzyme; ATP-proton motive force
 interconversion"
 /note="Residues 1 to 139 of 139 are 99.28 pct identical to
 residues 1 to 139 of 139 from *Escherichia coli* K-12 Strain
 MG1655: B3731"
 /codon_start=1
 /transl_table=11
 /product="membrane-bound ATP synthase, F1 sector,
 epsilon-subunit"
 /protein_id="AAG58934.1"
 /db_xref="gi:12518588"
 /translation="MAKTIYHLDVYSAEQMFSGLYEKIQVYSGSEGLCTPGAPLLT
 ATKPMQIRVYQGHGEETIYLSGGILEYQAGNVTYLAQTLRGODLDEARAMEKRRRA
 EKHSSHQDDVYQAQSAELAKKAIQLRVIELTKKAM"
 cds
 complement(2274. .3656)
 /gene="atpd"
 /note="55230"
 complement(2274. .3656)
 /gene="atpd"
 /function="enzyme; ATP-proton motive force
 interconversion"
 /note="Residues 1 to 460 of 460 are 100.00 pct identical
 to residues 1 to 460 of 460 from *Escherichia coli* K-12
 Strain MG1655: B3732"
 /codon_start=1
 /transl_table=11
 /product="membrane-bound ATP synthase, F1 sector,
 beta-subunit"
 /protein_id="AAG58935.1"
 /db_xref="gi:12518589"
 /translation="MATKTIYQVIGAVVDEFPDQAVPRVYDALEVQNGNERLYLEVQ
 QOLGGIVRTIAMGSDGLRGLDKLEHPTEVPVKGATLGRIMNVLGEYVDKRGCI
 GEEMNAHRAAPSYEELNSOELLEETGIKYIDLMCPAKGKGYLGGAQVGTVMK
 MELINILAIHSGSYFVAGVGEREGEDFHEMTDSVAVIDKVLAYVQNMPEPGNLL
 RVALGITMAEKFREGRDVLFPDNIYRTTLTLAGTEVSALGRNPSAVGTOPTIAEEMK
 GVALGRITSTKGTSTTSQAVAVYPADDTLDPSPATVAHLDAITVLSRQASLIGIYPA
 VDPDSTSRQDPLVVGQEHDTARGVOSIILQRYOEKLDITAIIGMDELSEEDLVVVA
 RARKIORTLSPDFVAEYFETGSPKYYVSLKDTINGKIMGEYDHLPEQAFYVWGSIT
 EEAVERKAKKL"

gene
 complement(3683. .4546)
 /gene="atpg"
 /note="55231"
 cds
 complement(3683. .4546)
 /gene="atpg"
 /function="enzyme; ATP-proton motive force
 interconversion"
 /note="Residues 1 to 287 of 287 are 100.00 pct identical
 to residues 1 to 287 of 287 from *Escherichia coli* K-12
 Strain MG1655: B3733"
 /codon_start=1
 /transl_table=11
 /product="membrane-bound ATP synthase, F1 sector,
 gamma-subunit"
 /protein_id="AAG58936.1"
 /db_xref="gi:12518590"
 /translation="MAGAKKEIRKSAVQNTOKITLAMEWVAASGMKRSODRMAASRP
 VAEIRAKYTIQHAGNLEFKRHPYLEDPRVKNQVLYSTDRGLCGKININIFKLLAB
 MKTITDKVQCDLAMIGSKGVSPRSVGGANVAQVYTGKNDPSLSEILIGPKWMLAY
 DEGRIDKLYIVSNKFINTMSQVPTISQLPLPASDDDLKHKSMWYIYEDPKNALDTE
 LILRRVEQYVQGVENIASQEAARMAVMKRAATNGSLLIKEQLVYNNKRAQASITDDE

gene
LITEIVSGAANV"
complement(4597. .6138)
/gene="atpA"
/note="5233"
complement(4597. .6138)
/gene="atpA"
/function="enzyme; ATP-proton motive force
interconversion"
/note="Residues 1 to 513 of 513 are 100.00 pct identical
to residues 1 to 513 of 513 from Escherichia coli K-12
Strain MG1655: B3734"
/codon_start=1
/transl_table=11
/product="membrane-bound ATP synthase, F1 sector,
alpha-subunit"
/protein_id="AAG58937.1"
/db_xref="GI:12518591"
/translation="MQLNSTELSELKRIQAFQVNVSEAHNEGVTSVSGRIHGL
ADCMQEMSLDEGNRTAATLNDERDSVGAIVMGYPADLABEMKVCYKGRILEPVYCG
ILGRVNTLGAIIDGKGPLDHDGFSVAEVAIPGVIERQSDVPQVGTAKVDSMTPIG
KQREELIIVDROGTGTALADAIDAINORDSGIKCIYVAGKASTISNVKLEHGLAL
NATVIVVAGSASALQOTLAPAGCAMEYERDGEDALIYDLSKQAAVAYRISLL
LRPRGREFPGCDVYLHSRIELERARVNAVYVAFPTGVEKGTSLTLPTEQGA
GVSAFVPIRNVSTIDGQIFLEFNLFNNGIRPAINPGISVSRGGAQOTIMKLTSG
ITFALAQVRELAFSQFASDLDLDTATRKOLDGQVTELLKQYAPMSVAQSQIVLRA
ARGVLADELKISGPSFPAALLAYVRDHAFLPQEIINOTGGVNDIEIGKLGIDSPK
ATQSW"
complement(6151. .6684)
/gene="atpH"
/note="25233"
complement(6151. .6684)
/gene="atpH"
/function="enzyme; ATP-proton motive force
interconversion"
/note="Residues 1 to 177 of 177 are 100.00 pct identical
to residues 1 to 177 of 177 from Escherichia coli K-12
Strain MG1655: B3735"
/codon_start=1
/transl_table=11
/product="membrane-bound ATP synthase, F1 sector,
delta-subunit"
/protein_id="AAG58938.1"
/db_xref="GI:12518592"
/translation="MSEFITYARPYAKAFDFAVEHOSVERMODMLFAAEVTKNEQM
ALLGALGAPETLASEPIFVACGEQIDENGQMLIFVMAENGRIHALPDLQFHLRAV
SAPTEEVQIVISAALSBOQLAKISAMEKRRLSRVYLCKIDKSVMGVILIRAGDWIT
DGSVAGRIELRADVIQS"
complement(6699. .7169)
/gene="atpF"
/note="5233"
complement(6699. .7169)
/gene="atpF"
/function="enzyme; ATP-proton motive force
interconversion"
/note="Residues 1 to 156 of 156 are 100.00 pct identical
to residues 1 to 156 of 156 from Escherichia coli K-12
Strain MG1655: B3736"
/codon_start=1
/transl_table=11
/product="membrane-bound ATP synthase, F0 sector, subunit
b"
/protein_id="AAG58939.1"
/db_xref="GI:12518593"
/translation="MNLNATITLGOAIAFVLFVFCMKYVMPPLMAAIEKROKELADCL
ASAEEASATDOLAKKAAEQVITIEQANRRSOTIDEAAEAEQETKIV
AOAQEIEARERAEELRKQVALILVAGAEKIEERSYDEKANSIDIVDKLVAL"
complement(7231. .7470)
/gene="atpE"
/note="5235"
complement(7231. .7470)
/gene="atpE"
/function="enzyme; ATP-proton motive force

VERSION AP002566.1 GI:13363930
KEYWORDS Escherichia coli O157:H7 (strain:O157:H7, sub_strain:RIMD 0509952)
SOURCE DNA.
ORGANISM Escherichia coli O157:H7
REFERENCE Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
AUTHORS Escherichia.
1 (sites)
Makino,K., Yokoyama,K., Kubota,Y., Yutsudo,C.H., Kimura,S.,
Kurokawa,K., Ishii,K., Hattori,M., Tatsuno,I., Abe,H., Iida,T.,
Yamamoto,K., Ohnishi,M., Hayashi,T., Yasunaga,T., Honda,T.,
Sasakawa,C. and Shinagawa,H.
Complete nucleotide sequence of the prophage VR2-Sakai carrying the
verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7
derived from the Sakai outbreak
Genes Genet. Syst. 74 (5), 227-239 (1999)
20198780
2 (sites)
Ohnishi,M., Murata,T., Nakayama,K., Kuhara,S., Hattori,M.,
Kurokawa,K., Yasunaga,T., Yokoyama,K., Makino,K., Shinagawa,H. and
Hayashi,T.
Comparative analysis of the whole set of rRNA operons between an
enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an
Escherichia coli K-12 strain MG1655
Syst. Appl. Microbiol. 23 (3), 315-324 (2000)
20557356
3 (sites)
Yokoyama,K., Makino,K., Kubota,Y., Watanabe,M., Kimura,S.,
Yutsudo,C.H., Kurokawa,K., Ishii,K., Hattori,M., Abe,H., Iida,T.,
Yamamoto,K., Hayashi,T., Yasunaga,T., Honda,T., Sasakawa,C. and
Shinagawa,H.
Complete nucleotide sequence of the prophage VR1-Sakai carrying the
Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli
O157:H7 strain derived from the Sakai outbreak
Gene 258 (1-2), 127-139 (2000)
20564182
4 (sites)
Hayashi,T., Makino,K., Ohnishi,M., Kurokawa,K., Ishii,K.,
Yokoyama,K., Han,C.-G., Ohtsubo,E., Nakayama,K., Murata,T.,
Tanaka,M., Tobe,T., Iida,T., Takami,H., Honda,T., Sasakawa,C.,
Ogasawara,N., Yasunaga,T., Kuhara,S., Shiba,T., Hattori,M. and
Shinagawa,H.
Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12
DNA Res. 8 (1), 11-22 (2001)
21156231
5 (bases 1 to 280900)
Ohnishi,M., Kurokawa,K., Makino,K., Yasunaga,T., Shinagawa,H. and
Hayashi,T.
Direct Submission
Submitted (26-JUN-2000) Ken Kurokawa, Osaka University, Genome
Information Research Center; 3-1, Yamadaoka, Suita, Osaka 565-0871,
Japan (E-mail:ken@gen-info.osaka-u.ac.jp,
URL:http://www.gen-info.osaka-u.ac.jp/, Tel:81-6-6879-8365,
Fax:81-6-6879-2047)
genome project.
location/Qualifiers
1. 280900
/organism="Escherichia coli O157:H7"
/strain="O157:H7"
/sub_strain="RIMD 0509952"
/db_xref="taxon:83334"
complement(143..616)
/gene="ECS4456"
complement(143..616)
/gene="ECS4456"
complement(143..616)
/note="similar to YIAI_ECOLI YIAI_ECOLI g1|2367245 percent
identity 98 in 157 aa (conserved in E.coli K-12),
/codon_start=1
/transl_table=11
/evidence="not_experimental"
/product="hypothetical protein"
/protein_id="BAB37879.1"

gene
CDS
/db_xref="GI:13363931"
/translation="NNREITADATKICGRCCEVACANSHOENDCAALSPDEFISRI
RVTKDSHTVTAACHQCEDAPCANVCPCVACSLREHGIIFVQTRICIGCKSCMLCPFG
AMEVVSRRKKARAIKCDLCWREHETGPACVEACPTKALQCMDEKVRHRLRQOPV"
complement(860..1675)
/gene="ECS4457"
complement(860..1675)
/note="probable transcriptional regulator, weakly similar
to hypothetical transcriptional regulator Y1SR [Bacillus
subtilis] g1|31233061|sp|P40331 percent identity 24 in 276
aa"
/codon_start=1
/transl_table=11
/evidence="not_experimental"
/product="putative transcriptional regulator"
/protein_id="BAB37880.1"
/db_xref="GI:13363932"
/translation="MLELSTLPTIVONGGLFISRGVGRHSARLESEIIFVEKRL
TIQERERFLVDAGSLILMPRORVGVEDPPADIKFTWLFYKAPRNDPRWLTHS
VPQQRVVDPAALVALFRQFUSEDEKHOETPALEFVLLIQOLTRDRALDEQAGAG
TTLAMKAOGLINTQYHRPLSTALARELYCNADYLGRRYRFFHLTLTEALHROVRE
AEKLISEARSLTEVAALCGFNDVGFQRIFRKHTGLTPAMKRRYSKEHNS"
1901..3301
/gene="ECS4458"
1901..3301
/note="probable permease, similar to hypothetical proteins
[Salmonella typhimurium] g1|7442781|p1|IC65167 percent
identity 37 in 444 aa, also similar to transport proteins
(putative symporters) e.g. Y1CJ [Escherichia coli (K-12)]
g1|2851421|sp|P31435|Y1CJ_EC"
/codon_start=1
/transl_table=11
/evidence="not_experimental"
/product="putative permease"
/protein_id="BAB37881.1"
/db_xref="GI:13363933"
/translation="WTSPTITPADIAKAKIDDAVSREKIGYGDAGGYITCLIM
FLTEFYTVFGLTALVGLTALNVPFALSDPVGVIAIDRQSMGFRFWMQIMTAI
PIGIIGITFLVFPDASMGVKIAMAFTVLLSVGTALNVYCARINTMTRHNEVIS
COSMREVLGVAGFLSVGLPMVMVALFEOGNARVGLGVGLCAIVAMFLCEPFVY
REVPDLSMGKFTLREHLAGLRNNDOLIMVMSLLINVRINRGAGVETVYVAG
STGYSLFEPVYVTPASITIGYSIVSRTPRPVYKLYLLINLAIAMETPLBSCGA
YQTLMAVILNGVILGFTLLPLHPSLAFADIDYGMKTHVSSGGRNPFNLFETLAN
ASSAGIISLLEFVAYQGVENQASSLGGIAMEETLLPALFHLLAGAIRPCKLNP
MMSRIATDLRQHRQP"
3312..5282
/gene="ECS4459"
3312..5282
/note="unknown, similar to hypothetical protein [Thermotoga
maritima] g1|7452109|p1|F72395 percent identity 37 in
635 aa"
/codon_start=1
/transl_table=11
/evidence="not_experimental"
/product="hypothetical protein"
/protein_id="BAB37882.1"
/db_xref="GI:13363934"
/translation="WNISEVDLHKLTIVSDPFLGOYQOALVRDVVIPYOMDALNDRIPEA
EPSHAIEFRFIAAGLOEGEFGYGMVPODDVAKWLEAVAMSCQKPDALERTAEVITE
LIASQCEDGLNTFTYKAPDERNSNLAECHELICAGHLIAGAVAFQATGKRRGLUG
VVCRLADHIDSVFGDESKLGYRQHPHPIELALMLKLVETEPRLATJTNFVORGA
OPHYDDOYERKRGQTSNHWITGYPAMVNDKAYSQLNLELAQOQTAIGHAVFVYLMGT
VAHLARLSHDSKRODCRLNMMNAOROLYITGGISGSSGEAFSDVDLPNDVYAA
SCASIGLMFARBMILEMGDSOYADVMEBALYVNYLGGMALDGKHFVNVLEWPKS
LKNPIYDHYVPIROHWRGCAACCPNTARVTFSGHYTVYREDAITNINAYAGSMVY
PVENGSTLRVSGNPQMOGYTTAVESQVPRHRLTALPLPMTOPQITLLNGEYBOD
IRKGLYITREMGEEDTLNLTLPMEVRRVRYVAVRHAGVVALIQRGFLVYCLEADN
GESLHNLMPLPADAPFTTEGKGLFSKLLIDAPGRYDOSNPBODPLMHWDSABAKRO
TOTLTFIPWFSNARGEMERIWNEEKHP"
complement(5412..6152)

D	223723	CGATTTCCCTCAGGATGTCGCGACCGCGGTGTAAGATGCTTTGAGGTGCAAAATGTTAA	223664
Q	1349	aacaggtttgtagtctgaagttcaacaacaatgaigtgtgtgtatgttcgtgtatcgc	1408
D	223663	TGACCGCTCTGTGTGTGTGTAAGTTTCAAGCGACGCTCGGGGCGGTATCGTGCACATCGC	223604
Q	1409	aattgagatcatctgattgatttaaacacgcgatttaagcgttaacaataatcgaataccaat	1468
D	223603	AATGGTTTCCTCCGACCGCTGTCGCTCCGGCTTCGTGATGTAAAAGACCTCGAACACCCGAT	223544
Q	1469	tctgttctcagttggaacggaacaatltggttcgtatcatatgaacgtlatgttgtaaccaat	1528
D	223543	CGAAGTCCCCGGTAGGTAAACGACACTCGGCCCGATCATGAAGTACTGAGGTGACACGGT	223484
Q	1529	cgaatgagcaaaagtgaaatcgtgttcgagaaagaaattggtctattccacgcygcgcacaag	1588
D	223483	CGAATATAAAGGCGAGATCGGTGTAAGAAAGCGCTTGGGCGATTCCCGCGACACCTTTC	223424
Q	1589	ttaatgaagaacaactacacagctactgaactttagaacacggaatlaaagtltacact	1648
D	223423	CTACGAGAGACCTGTCAAACTCTCAGGAACACGTGCGGAACCGGTATCAAAATTATCAGCT	223364
Q	1649	agttgttcggttttggaaggggtlaaagttatctcgtgtgtgtcggtgttcggtlaa	1708
D	223363	GATGTGTCCTCCCTCAAGGCGGTAAAGTTGGTCTCTGTTGGTGTGCGCGGTAGGTAA	223304
Q	1709	aaccgttaatatgttgtaataatccgtlaaacatcgcaattgaacacatcaggttaactgt	1768
D	223303	AACCGTAAACATGTATGGAAGCTCATCTGAACATCGGCATGAGCACACTCGGTTACTGTG	223244
Q	1769	ctttgcgggggttagtgtagctacgcgtgaaaglaaacagactctatcatatagaatgaaga	1828
D	223243	GTTTGGCGGGGTAGTGAAGTACTCGTGAAGGGGAAGCACTTCTACACAGAAATGACCGA	223184
Q	1829	ctctaaacgtattagataaagtgtctctctgttttatgttcaaatgaacgagccaccagtaa	1888
D	223183	CTCCAAAGTTATGCATTAATATATATCCCTGGGTATGGCCAAATATAGAGCCGCCGGGAAA	223124
Q	1889	ccgttgcgtgtgcatctaacagcgttaactatcgtgcgaaanaaatccgtatgaagtctg	1948
D	223123	CCGCTTCGGGGTAGCTCTTGACCGGCTCGACACATGCGTGAATAAATTCGCTGACGAAGTGC	223064
Q	1949	tgatgtcttatcttcgtgtgataa	1972
D	223063	TGAGCTTCTGCTGTTCCGTTGACAAA	223040

FEATURES	source
LOCUS	AF188265.1
DEFINITION	AF188265.1 7421 bp mRNA
ACCESSION	Salmonella typhimurium strain 7421, ATP synthase operon, complete sequence.
VERSION	AF188265
KEYWORDS	AF188265.1 GI:6625695
SOURCE	
ORGANISM	Salmonella typhimurium. Salmonella typhimurium Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Salmonella.
REFERENCE	1 (bases 1 to 7421) Kim, H.-K., Heo, N.-J., Ghim, S.-Y. and Song, B.-H.
AUTHORS	
TITLE	Molecular structure of nonastrictonic atp genes encoding ATP synthase in Salmonella typhimurium
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 7421) Kim, H.-K., Heo, N.-J., Ghim, S.-Y. and Song, B.-H.
AUTHORS	
TITLE	Direct Submissions
JOURNAL	Submitted (24-SEP-1999) Dep. of Biology Education, Kyungpook National University, #1370 Sankyuk-dong, Buk-Ku, Taegu 702-701, Korea
FEATURES	Location/Qualifiers
source	1..7421 /organism="Salmonella typhimurium"

gene . /strain="Ta98"
328. .7227 /db_xref="taxon:602"
gene /gene="ATP synthase operon"
328. .720 /note="atpI: uses putative GTG translational start site"
CDS 328. .720 /codon_start=1
/transl_table=11
/product="ATP synthase subunit 1"
/protein_id="AAF19355.1"
/db_xref="GI:6625696"
/translation="MKNWSVSLVSRNVARLLLVOLLVITASGLFSLKDPKCVSA
ISGGLAVLPVNLPMIFAMRHOAHTPAAGRAVMPFAGEAFKVLAMLVLLVALAVLK
AVPLIYTWVVLVAVQILLAPAVINNGK"
729. .1544 /gene="atpB"
729. .1544 /gene="atpB"
CDS 729. .1544 /note="atpB"
/codon_start=1
/transl_table=11
/product="ATP synthase subunit a"
/protein_id="AAF19356.1"
/db_xref="GI:6625697"
/translation="MASENMTPODYIGHLNNLQDLRTFSLVDPONPAPATFWINID
SMFESVVLGILFIVFRSVAKKATGYGKFORAPILVIGFVNGSVKDMYGRKSLA
PLALTFVWVLAMLMDDLPTDLPLTYAENHVGIPALRVVPSADVNTLSMALGVFLIL
ILFYSIKKGGIGGFKETLTQPFNMMAFIPVNLILEGVSLSKFPVSGLRFLFGMAG
ELIFLILGLPLPMSOWILNVPMAFILIIITLQAFIMALTIVYLSMASEH"
1591. .1830 /gene="atpE"
1591. .1830 /gene="atpE"
CDS 1591. .1830 /note="atpE"
/codon_start=1
/transl_table=11
/product="ATP synthase subunit c"
/protein_id="AAF19357.1"
/db_xref="GI:6625698"
/translation="MENLMDLYMAAAVMMGLAIGAIGILGSKFLEGAARPPD
LILPLRTQFIFVGMGLVDAIPMAVGLGYMFAVA"
1892. .2362 /gene="atpF"
1892. .2362 /gene="atpF"
CDS 1892. .2362 /note="atpF: uses putative GTG translational start site"
/codon_start=1
/transl_table=11
/product="ATP synthase subunit b"
/protein_id="AAF19358.1"
/db_xref="GI:6625699"
/translation="MNLNATILGOAIVLFLFCMKVYMPPLMAAIEKROKIANGL
ASAEKHAHDLDLAKASATDQKKAEQVYIEQANRRSQIIDEAFAEEQOEFTKIV
AQAQIEIAEKKRRAREELRKQVAILAVAGAEKIERSVDERANDSDIVDKIVAEI"
2377. .2910 /gene="atpH"
2377. .2910 /gene="atpH"
CDS 2377. .2910 /note="atpH"
/codon_start=1
/transl_table=11
/product="ATP synthase delta subunit"
/protein_id="AAF19359.1"
/db_xref="GI:6625700"
/translation="MSERTVAPRYAKAFDFAVEHOSVERMODPLAFAAEVTYKNEOM
AELISGALAPETLIESFIVAGBOLDENGOUILVMAENGLNALPVLPOEFLRAVY
SEATRENVITAAALSTEOQLAIEISAAEKKRISRVKLNCKIDKSVMAGYIIRAGDWIY
DGSVGRERIEADVIQS"
2923. .4464 /gene="atpA"
2923. .4464 /gene="atpA"

CDS
2923..4464
/gene="atpa"
/note="atpa"
/codon_start=1
/transl_table=11
/product="The ATP synthase subunit alpha"
/protein_id="AAFI9360.1"
/db_xref="GI:6625701"
/translation="MOLNSTEISELVKQRIAPNVNVEANNEGTVISVSDGVIIRHGL
ACMGEMISLPGNRAIALNLEBDSVAVVAGPYADLAEGMKVKCTGRILIEPVVGR
LLGRVNTGAPIDGKPLDHGFSAVEAIPGVLEKROSDVDPQVOTKAVDSMIG
RQURELITDROTKTALAIDAIINRDSGKICITVATOKASTISNVVRIKEEHAL
ANTIVVATASSEALQYLAIPYACAMGEYFRDGEDALLITDDSKQAVATQISLL
LRPPGRAPGVDPVLEHSLRLERARVNAVEATKGEVGRGSLAPLPIETQA
GVSAVPTNPVISTIDGQIFLETNLENAIGIRPVPNGISVSRVGAAGQKIMKLSIG
ITPAIOLYRELAIFSQFASDLDTAKRODHGKVELLKOKOYAPMSVAGOSLVFA
AERGIADYVELKISFEAALLAYVDRDHAFLMQLINQGVNDELKGLKGLDSFK
ATDSW"
/gene="atpg"
4515..5378
/gene="atpg"
4515..5378
/note="atpg"
/codon_start=1
/transl_table=11
/product="ATP synthase subunit gamma"
/protein_id="AAFI9361.1"
/db_xref="GI:6625702"
/translation="MAGAKETRSKIASVONTOKITKAMENVAASKRRKSODRMAASP
YATMKRVIGHLAHGNLEKHPLEDRDVRVGVLYVSDRGICGGINILTFKLLAE
MKVTDKGVQCDLAMIAGSKVSFFDSVGVNVAQVGMGNPESLIEIGVYKWLQAY
DEGRDLKLYIVSKNFINTMSQVPTISQLPLPASDDDLKHKMDVILEDPKALDT
LIRVYEQVYQGVENVLASEQAARVMAKATDNGSLIKELQVYNKARASTIOE
LIRIVGAAY"
5405..6787
/gene="atpd"
5405..6787
/gene="atpd"
/note="atpd"
/codon_start=1
/transl_table=11
/product="ATP synthase subunit beta"
/protein_id="AAFI9362.1"
/db_xref="GI:6625703"
/translation="MATGKIVQVIGAVVDEFPDAPVRYDALELVONGNERLVLEVO
QOIGGIVPTIAMSSDGLRGDLVDKLEHPTEVPVKATLGRIMNLGEVDMKGEI
GEEBRRAIHRAPSYEELNSOELLETFIVIDIMCPAPAGKYGGLGAGVGTVM
MELIRNTAIESGTSVAGCBERTRENDYTHENTDSNVIDKYSIYVGMNEPGRNL
RVALGTITMAEKFEDEGRDYLLEVDNLYRTTLAGEVSALGRMPSAVGTOTLAEM
GVLOERTSTRTGSLTSVQAVVYPADLTPSPATTAHLDATVLSRQISLSTIPA
VDPIDSTRQDLPLVGOEHYDPAQVOSILORQELKDVIAIGMDLSEEDLVVA
RARKIORLSOPFEVAVETGSPKYSVLKDTINGFKIMGEVDHLPEQAFYVWGS
EEAVERAKKL"
6814..7227
/gene="atpc"
6814..7227
/gene="atpc"
/note="atpc"
/codon_start=1
/transl_table=11
/product="ATP synthase subunit epsilon"
/protein_id="AAFI9363.1"
/db_xref="GI:6625704"
/translation="MTYHLDVYSAEQMFSGLVKRIQVTSSEGEIGTIPGHALLTAT
KPGAIRIVKQGHHEFTYLSGILEENQGRNTVLAADPIAQQDDDEARAMAKKAAEE
HISSHGVDDAQAASAEIAKAIADLRVIELTKKAM"
BASE COUNT 1764 a 1854 c 2020 g 1783 t
ORIGIN

Query Match 41.0%; Score 808.8; DB 2; Length 7421;
Best Local Similarity 64.6%; Pred. No. 2.3e-148;
Matches 1294; Conservative 0; Mismatches 662; Indels 48; Gaps 4;

OY 11 ttggctcaagtctggcttgatctctatgatgacacagtaagaacattatcagatgca 70
DB 4138 TTTCTTCAGTTTGCATCCGACCTTGACAGATGCAACTCGTAAGACGTTGACCGGTTCA 4197
OY 71 aaagtaactaatattatgaacaaacaaactactctccgttactctgtaagacaaagca 130
DB 4198 AAGTACACCAATGCTGTAACAGAAACAACTATGCGCGATGTCGTTGGCAGCAGTCT 4257
OY 131 ctatgtaattctcagtagaagtttgattacttaagaagaatgtaactagatctgtt 190
DB 4258 CTGCTTCTGTTCCAGCAGCAACGCTGTACCTGCGGATGTTAACTGTGAAATCGGC 4317
OY 191 tcaattgaatcagcactttaagtagtcaatacactaactctgatttgaatgctgag 250
DB 4318 AGCTTCAACCCGCTGCTGCTGCTTACGTGACCGGTGATACCGTCCGTTGATGCAAG 4377
OY 251 ttaaccatcttgcaatcttaaatatgaatgaatgaatgaatgaatgaatgaatgaat 310
DB 4378 ATCAACCAACGCGGTGCTTACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4437
OY 311 agctcaagaacaaacagtgctgta-----agtaacact 346
DB 4438 TCTTCAAGCAACCAATCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4497
OY 347 ttaattgagagacaaatgagtgagtgtaagaagaatgaagaacaaatcgagtgtaa 406
DB 4498 GCATTTAGAGAGAACCTCATATGCGCGGCGCAAAAGATACCTAGTAATGCAAGCCTCC 4557
OY 407 aaagtaacaaataacttaacaaacagatgtaaaatgctgctcgtgaaatgctgtaaa 466
DB 4558 AGAACACGCAAAATATCATCTAAACGATGAGATGTCGCGCTTCCAAATGCGTAAAT 4617
OY 467 cgcagaacgcagtgctctctcagcccttctaagaacaaatcgttaagtggtggcc 526
DB 4618 CGCAGATCGCATGCGCGCCGCGCGCTCTTATGCAAAACATGCGCAAAATGCTGCTC 4677
OY 527 acgttccaagaacagatggttgaacaaatcatttgaatgagtcgcaagtaaaa 586
DB 4678 ACCTTGACACGATGATGTAATGTAATGACACCTTACTGGAAGACCGCATGTAAC 4737
OY 587 aagtgagcatgattgctgctcacaagatcgctggtcttctgctgctgctgctgctgct 646
DB 4738 GCGTGGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4797
OY 647 tcttaaaactgtaataaagaaatgaagaagaagaagaagaagaagaagaagaagaaga 706
DB 4798 TGTTCAAAACCTGCTGCGGCAAAATGAAGACCTGACCGCAAAAGCGCTTCAATGCGACC 4857
OY 707 tgaatttaactggttcttaactatcaactcttcccaatcttgggaattaaatttaa 766
DB 4858 TCGCAATGATGCGCTCGAAAGCGGTGCTTCTGCACTCCGTGGCGGCAATGTTTGG 4917
OY 767 cccaagattcagtagtgatatactccctcgttgaagcaglttaattggttcagtaatt 826
DB 4918 CCCAGTACCGGCTGAGGATTAACCTTCCGCTGCGCAACTGATGCTCGGTAAG 4977
OY 827 ctatgattgattctataaaagaagggaagtagatggttgattagtttaatacaaat 886
DB 4978 TGATTTGACGCTTACGACGAAAGCCGCTGACAACTTACATTTGTCAGCAAAAT 5037
OY 887 ttataacacgattcgcaaaaagccagcatggaataataatccattaccagaattag 946
DB 5038 TTATTAACACCACTGCTCAGAGTTCCGACATCAGCCACTGCTGCTGCTGCTGCTGCT 5097
OY 947 ataagtaagaattagcgaaagaacaaagtttggaattatatactgaacacgtatgca 1006
DB 5098 ATGATGATGATCT-----GAACAATTAATCTGAGTATTAACGACCAACCAATCGCA 5151
OY 1007 aagtaattatgaataattattggttgcttattgaatctcgaagttatcaagaacagcag 1066
DB 5152 AGCGGTGCTGATACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5211

1067 ttgaaacctgtcttctagcaagccgcgtcgaatgtgtccatgaagaagcaagaata 1126
 1127 acgcagtaacttaataatgagttacagttatgtatatacaaaactcgtcaagaaga 1186
 5272 ATGGGCGACGCGTATTAAAGAGTTCAGATTGATACAAACAAGCTGTCAGGCGACGA 5331
 1187 ttcaaatgaatataatgaatgtgtcgcgtgtcgaagcaaatgaaca-----a 1237
 5332 TTACTCAGGAACCTACCGAGATCGTCTGGGGGCGCGCGGCTTTAAACAGGTTATTTCG 5391
 1238 tagaagatcgttaatgagcaactggaataatgttacaatacactcgtgtgattagact 1297
 5392 TAGAGGATTTAAGATGCTACTCTGGAAGATTGTCCAAGTAATCGCGCGCTGATGAGCT 5451
 1298 tgaatcccaacaagaatgaatgaatcaaaagatataatgact-----aaagtga 1348
 5452 CGAATTCCCTCAGAGATGCGGTACCGCGGCGGTACGATGCTGTGAGGTCGAATAATGTTAA 5511
 1349 aacaggttagtacttgaagtcaacaacaataatgagttgtgtgtgtgtgtgtgtatgc 1408
 5512 TGAGCGTGTGGTGGTGAAGTTTACACACAGCTCGCGCGGTATCGTCCATCCATCCG 5571
 1409 aatggatcactgtatgattaaacgcggtttaagcgttaacaataatgaataaccact 1468
 5572 AATGGGTTCTTCGCGGCTCGTCCGCGGTGTGATGTTAAAGACCTCGAACACCCGAT 5631
 1469 ttctgttcagttggaagcaaaacattgtgtgtatcatgaacgttatgtgtgaaccaat 1528
 5632 CGAATGCCCGGAGTAAAGCATCTGCGCGGTATGATGATGATGATGATGATGATGATGAT 5691
 1529 cgaatgaagaagttgaatcgtgtcgtcgaagaagaattgtgtatcatcaccgtgtgcacaa 1588
 5692 CGACATGGAAGCGGAGATGCGGTGAAGAGCGGTGCGGTGATGATGATGATGATGATGATGAT 5751
 1589 ttatgaagaacaatcaacagtaactttagaagcgttaataatgaataatgaact 1648
 5752 CTACGAAGAAGCTGTAAACTCTCAGGAACCTGCTGGAACCGGTATCAAAAGTATGACACT 5811
 1649 agttgttcgtttcgaagaaggggttaaaagtaagttatcgtgtgtgtgtgtgtgtgtgt 1708
 5812 GATGTGTCGTCGTTAGAGGCGGTAAGATTGCTGTGTCGTCGTCGTCGTCGTCGTCGTCGTC 5871
 1709 aacgtcaatatgattgagtaataatcgcgaataatgaacattgaacattgaactgt 1768
 5872 AACCTAAACATGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5931
 1769 ctgtgaggggtgagttgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgagcgtgag 1828
 5932 GTTTCGGGGCGTGAAGTGAACGTAACGTAACGTAACGTAACGTAACGTAACGTAACGTAACGTA 5991
 1829 ctctaactatataagtaagatgtctctgtttatgttacaagaagcgcacacagttaa 1888
 5992 CTCACAGCTTATGCAAAAGTATCCCTGCTGATGATGATGATGATGATGATGATGATGATGAT 6051
 1889 ccgtttacgtgtgagcattacaagcgttactactatgagcgtgagcgtgagcgtgagcgtgag 1948
 6052 CCGTCTGCGGCTGCTGCTGAGCGGCTGACCATGCTGAGAAATTCGTTGAGCAAGGTCG 6111
 1949 tgatgtctatcttctgtgtataa 1972
 6112 TGACGTTCTGCTGCTGTTGACAA 6135

SOURCE
 ORGANISM
 Escherichia coli DNA.
 Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 REFERENCE
 1 (bases 3160 to 5694)
 Kanazawa, H., Kayano, T., Mabuchi, K. and Futai, M.
 Nucleotide sequence of the genes coding for alpha, beta and gamma
 subunits of the proton-translocating ATPase of Escherichia coli
 Biochem. Biophys. Res. Commun. 103, 604-612 (1981)
 2 (bases 506 to 2622)
 Kanazawa, H., Mabuchi, K., Kayano, T., Nouni, T., Sekiya, T. and
 Futai, M.
 Nucleotide sequence of the genes for F-0 components of the
 proton-translocating ATPase from Escherichia coli: Prediction of
 the primary structure of F-0 subunits
 Biochem. Biophys. Res. Commun. 103, 603-620 (1981)
 3 (bases 812 to 3311)
 Nielsen, J., Hansen, F.G., Hoppe, J., Friedl, P. and Von Meyenburg, K.
 The nucleotide sequence of the atp genes coding for the F-0
 subunits a, b, c and the F-1 subunit delta of the membrane bound
 ATP synthase of Escherichia coli
 Mol. Gen. Genet. 186, 333-339 (1981)
 4 (bases 3169 to 4701)
 Gay, N.J. and Walker, J.E.
 The atp operon: Nucleotide sequence of the region encoding the
 alpha-subunit of Escherichia coli ATP synthase
 Nucleic Acids Res. 9, 2187-2194 (1981)
 5 (bases 389 to 3173)
 Gay, N.J. and Walker, J.E.
 The atp operon: Nucleotide sequence of the promoter and the genes
 for the membrane proteins, and the delta subunit of Escherichia
 coli ATP-synthase
 Nucleic Acids Res. 9, 3919-3926 (1981)
 6 (bases 5641 to 7540)
 Kanazawa, H., Kayano, T., Kiyasu, T. and Futai, M.
 Nucleotide sequence of the genes for beta and epsilon subunits of
 proton-translocating ATPase from Escherichia coli
 Biochem. Biophys. Res. Commun. 105, 1257-1264 (1982)
 7 (bases 1 to 817)
 Kanazawa, H., Mabuchi, K. and Futai, M.
 Nucleotide sequence of the promoter region of the gene cluster for
 proton-translocating ATPase from Escherichia coli and
 identification of the active promoter
 Biochem. Biophys. Res. Commun. 107, 568-575 (1982)
 8 (bases 2129 to 2599)
 Jans, D.A., Fimmel, A.L., Hatch, L., Gibson, F. and Cox, G.B.
 An additional acidic residue in the membrane portion of the
 b-subunit of the energy-transducing adenosine triphosphatase of
 Escherichia coli affects both assembly and function
 Biochem. J. 221, 43-51 (1984)
 9 (bases 1 to 7881)
 Walker, J.E., Saraste, M. and Gay, N.J.
 Nucleotide sequence, regulation and structure of ATP-synthase
 Biochim. Biophys. Acta 768, 164-200 (1984)
 10 (bases 565 to 957)
 Gay, N.J.
 Construction and characterization of an Escherichia coli strain
 with an uncl mutation
 J. Bacteriol. 158, 820-825 (1984)
 [9] revises other sequences.
 The atp synthase operon consists of nine genes coding for nine
 structural proteins of ATPase. It is not known through these

atp-1		uncI polypeptide		papD		atpB		uncB.
atp-2		a subunit		papH		atpC		uncC.
atp-3		c subunit		papF		atpF		uncF.
atp-4		b subunit						
atp-5		delta subunit		papE		atpH		
atp-6		alpha subunit		papA		atpA		unca.
atp-7		gamma-subunit		papG		atpG		
atp-8		beta subunit		papB		atpD		uncD.
atp-9		epsilon subunit		pape		atpC	[2] is	

revised by [7]; due to the large number of revisions, they have not been presented as individual revisions in the FATHOMS table.

		/transl_table=11	
		/protein_id="AA224733.1"	
		/db_xref="GI:148135"	
		/translation="MNLNTLTIGQATAFVLFVLFCMKYVWPLMAIEKROKEIADGL ASERKHLDLAKASATDQLKKAKEAOYLIDOKARRSQIIDENKAHEDEORETIV AQAQAEIEARERARELRKOVAILVAAGEKIITERSVDPAANSIDYDLKLVEL"	
gene		2672..3205	
		/gene="unch"	
CDS		2672..3205	
		/note="ATP synthase delta subunit (atp-5)"	
		/codon_start=1	
		/transl_table=11	
		/protein_id="AA24734.1"	
		/db_xref="GI:148136"	
		/translation="MSEFTIVARPYAARFAEDVEHOSVERMODLAFAEVTNRKDM HELLSGALAEETLAEFTIAVGGEOLDENGONLIRVAENGRULAMPDYLEOFLHTAV SEPTAEVDYISAALISEQQOLAKISAMERKLSRKVLNCKIDKSVMAGVITIRAGDWVI DGSVRGRLETRIALVLOS"	
gene		3218..4759	
		/gene="uncg"	
CDS		3218..4759	
		/note="ATP synthase alpha subunit (atp-6)"	
		/codon_start=1	
		/transl_table=11	
		/protein_id="AA24735.1"	
		/db_xref="GI:148137"	
		/translation="MOLNSTEISELIKORIAOFNVNSEANNEGTVSVDGVRIRHCL ADMCGEMISLPGNRYAIALLNERDSVGAAYMGPYDALAGMKVKCTGRILEVPVGG LCRWVTLGAPIDRGKPLDHDFSAVEALAIPGVTEROSVDOPVOTGKRVLEDPG RGREILIIIGROTGTALIDAIIINODSGIKTYVAILGOKASTISVWRKLEHGHAL ANTIVYATPSSESAAIOYLARMVALMGEXFERDGEALAIYDLSASAAYRIISLLA LNRRPGEARPPGDVFYTHSLRLERAAYVNVAEYDAFPKGVCKGTSGLTLPIIEQOA GDVASAFPYNVISITTDGQIFLETMLFENAGIRPAYNPDISYSRVGGAAQTIMKRLSSG IRRALAYPEIRLAAFSOPASDLUDATRNDQHGDGVTELLKQRTAPRSVAQOSIVLER AESRYLADVLSLKISGFEEAALLAYVDRDHAPLMQEIOMTGGYNDELIGKLKJLIDSPK ATWSM"	
gene		4810..5673	
		/gene="uncg"	
CDS		4810..5673	
		/gene="uncg"	
		/note="ATP synthase gamma subunit (atp-7)"	
		/codon_start=1	
		/transl_table=11	
		/protein_id="AA24736.1"	
		/db_xref="GI:148138"	
		/translation="MAGAKDIRSKIASVONTQKITKMENVYASNMKRSQDRMAASRP /ATWSM"	
Query Match	41.0%:	Score 808.8; DB 2; Length 7881;	
Best Local Similarity	64.6%:	Pred. No. 2,3e+148;	
Matches 1294; Conservative	0; Mismatches .66z; Indels 48; Gaps 4;		
Oy	11	ttaggtcagttcggttcggattcttaigtatgcacagcytaaagaatlatcacatggtcaa	70
Dd	4433	TTCCTCAGATTTCAGATCCGACCTTCAGATGCAGAACACGTAACCACTTGACCACGGTCCAG	4492
Oy	71	aagaactgaataattgaacaacaaccaatactcttcogtattctgltgacgaacaaga	130
Dd	4493	AAAGTAGCCGAACTCCTGTAACAAGAACAGTGTGGCGCATGTCGGTGGCAGCAGTGT	4552
Oy	131	ttagtgattataggagttagatttggttaacctagaagaagtgaacttatgatcgtattg	190
Dd	4553	CTGGTTCTGTTGGCAGCAACGCTGTTTACTTGGCGGAGTTCGAACCTGCCAAATTTGGC	4612
Oy	191	tcaattgaatcagcaactttagagtagtctaaccataactatcgtattttaatgctg	250
Dd	4613	AGCTTCGAAGCCGCTCTGCTGGCTTAAGTCGACCGCGATCACGCTCCGTTGATGCAAGAG	4672
Oy	251	ttaaccacaatctggaatatcacatgatgaataaagaagtagcatataaagcatttgt	310
Dd	4673	ATCAACCAACCGGTGGCTACAAAGCAAGAAATCGAAGCGCAAGCTGAAAAGGCAATCTCTGAT	4732

QY 311 agctcaagaacacagtcgtgta-----agtaact 346
 Db 4733 TCCTTCAAGCAACCAACCTCGTAACTCTGGCGGCTTGCCTTAGGCGAGCCCAAG 4792
 QY 347 ttaaatgagagacaaatgcaagtcgtctaaagaaagaaacaaatcgagtgta 406
 Db 4793 GCATTGAGAGAGAGTCAATGCGCGCGCAAAAGACATACATAGTAGTAGTCAAGGCTCC 4852
 QY 407 aagatcaacaaataataactaaagatgaaatggttcgtcctcgaataatgcytaaa 466
 Db 4853 AACAACGCAAGAGATCACTAAAGGATGAGATGCTCCCTTCCAAATGCGTAAT 4912
 QY 467 ccaagaacgcatgtcttcacacccatttcagaacaaataagtaagtaagtc 526
 Db 4913 CGCAGAGATGCAATGGCGGCGCACCGCTTATGCAAAACATGCGCAAGTATGGTTC 4972
 QY 527 acggttccaagaacagatggttcaacagatcattttatgtagtcggaagtaaaa 586
 Db 4973 ACCTTGCAACCGTATCTGGAATATAGCACCTTACCTGGAAGACGCGAGCTTAAAC 5032
 QY 587 aagtggaatggttctgttcacagatcggtcgttcgttcgttcgaatgtaact 646
 Db 5033 GCGTGGGCTACTGCGGTGCTGACCGAGCTGCTGTTGCGGCGGTTGACATTAAC 5092
 QY 647 tgttcaaaactgtatataatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 706
 Db 5093 TGTTCAAAAAATGCTGGCGGGAATGAACCTGACCGCAAAAGCGTTCAATGGGAC 5152
 QY 707 tgaattcaatcggttcaatcattcaatcatttccaaatcatttggaaatataat 766
 Db 5153 TCGCATGATCGCTGGAAGAGCGCTGCTGCTCAACTCGGCGGCGCAATGTTGTTG 5212
 QY 767 ccaagatcagatggttctgttcacacacacacacacacacacacacacacacac 826
 Db 5213 CCCAGGTACCGGATGGGGGATTAACCTTCCCTGCGAATGATGATGCGGTAAG 5272
 QY 827 ctatgattgattctataaaaggagagtagatggttgcattatgattataacaaat 886
 Db 5273 TCATGTGGAGCGCTACGAGCAAGGCGCTGACAGGTTTCAATGTCAGCAACAAT 5332
 QY 887 ttaataacagatgctgcgaagaacagatggaataatataatcattaccagaatg 946
 Db 5333 TTATTAACACCATGCTCTGAGTTCCGACCATCAGCAGCTGCTGCGTATACGCGCATCAG 5392
 QY 947 ataagatgaatagagcgaagaagaagaagttggattataattacagactgtagca 1006
 Db 5393 ATGATGATGATCT-----GAACAATAATCCTGGGATTAACCTGTAACCAACCGATCCA 5446
 QY 1007 aagatataatataattatggttcgttataatgaatcaggttataacagacagag 1066
 Db 5447 AGGCGTGTGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5506
 QY 1067 ttgaaacactgcttcgaagcagcgcgtcgatggttcgcatgaaagcagaacagata 1126
 Db 5507 TTGAAACACTGCGCCGCGAGCGCGCGCGCTGATGCTGATGCTGATGCTGATGCTG 5566
 QY 1127 acgagatgaatataatgattagttacagttatataacaaagctcgtcagaagta 1186
 Db 5567 ATGGGCGAGCTGATTAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5626
 QY 1187 ttaacaatgaatataatgattagttcgtcgtgagcagcaatataacaa-----a 1237
 Db 5627 TTACTACAGGAACCTACCGAGATCGTCTGGGGGCGCGCGGTTTAAACAGTTATTTCG 5686
 QY 1238 tagagatcgatataatgcaatgaaatgtaacaaatcaggttcggttattgagct 1297
 Db 5687 TAGAGATTTAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5746
 QY 1298 tgaattccacaaagatgagtaacaaagatataatgctt-----aaagtga 1348
 Db 5747 CGAATTTCCCTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5806
 QY 1349 aacaggtttagtactgagttcaacaaataggtggtggtgagttcgtatcgc 1408

Db 5807 TGAAGCTCTGCTGCTGGAAGTTAGCAGACAGCTGCGCGGCTATGCTAGCTACCTCC 5866
 QY 1409 aatggatcatctgatalgattaaacgcggtttaaagcgtataacaaataacaaat 1468
 Db 5867 AATGGTCTCTCCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5926
 QY 1469 ttcgttcagtgaggaaacaaatcattggttcgtatcagaaagatggttggtaaccaat 1528
 Db 5927 TGAAGTCCCGTAGGTAAAGCGACTGCGCGCTGATCAAGAGTACTGGTGAACCGGT 5986
 QY 1529 cgaatgagcagatgaaatcggttcgaagaagaatggttcattaccggtgcacaaag 1588
 Db 5987 CGACATGAAGAGGAGATATGCTGAAGAGAGCTTGGCGATTCACCGGACACTTC 6046
 QY 1589 ttatgaagaacaaactcaactagtaacttttagaagcgggaataagtaactgact 1648
 Db 6047 CTACGAAGAGCTGCTCAACTCTCAGGAGACTGCTGAACCGGATCAAGATTATCGACCT 6106
 QY 1649 agtttcggttcggaaggggtaagtagttatttcggtggtggtggtggtgta 1708
 Db 6107 GATGTCCTCGTCTGCTGAAGGCGGTAAAGTGTGCTGCTGCTGCTGCTGCTGCTG 6166
 QY 1709 aaccgcaatagatgaaatcaatccgtaacatcgcaatgagcaactcaagttactgt 1768
 Db 6167 AACCGTAACATGATGAGTACTATTTGTAACATCGGATCGACATCCGCTTACTCTGT 6226
 QY 1769 ctltcggggtagtgagcgtgagcgtgaaagtaagcaactcatalcatgataaga 1828
 Db 6227 GTTTCGGCGGTAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6286
 QY 1829 cttaacgattatagaaatggttcgttattatgtaataagaaagcaacacagta 1888
 Db 6287 CTCAACGTTATGACAAAGATATCCTGCTGATGCTGCTGCTGCTGCTGCTGCTG 6346
 QY 1889 ccgttaacgtgagcattacagcgttataactatgagcgaataatccgtgagtaagtcg 1948
 Db 6347 CCGTCTGCGCGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6406
 QY 1949 tgatgcttattcttcgtgataa 1972
 Db 6407 TGACGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6430

RESULT 12
 VAUNC 7402 bp DNA BCT 10-FEB-1999
 DEFINITION *Vibrio alginolyticus* unc operon for ATP synthase (EC 3.6.1.3).
 ACCESSION X16050
 VERSION X16050.1 GI:48331
 KEYWORDS ATP operon; ATP synthase; ATP synthase-coupling factor;
 DCD-binding protein; unc operon;
 SOURCE
 ORGANISM
 Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.
 REFERENCE
 1 (bases 1 to 7402)
 AUTHORS Krumholz, L.R.
 TITLE Direct Submission
 JOURNAL Submitted (09-AUG-1989) Krumholz L.R., Stanford University, Dept of
 Biological Sciences, Stanford CA 94305, U S A
 2 (bases 1 to 7402)
 Krumholz, L.R., Esser, U. and Simoni, R.D.
 TITLE Nucleotide sequence of the unc operon of *Vibrio alginolyticus*
 JOURNAL Nucleic Acids Res. 17 (19), 7993-7994 (1989)
 MEDLINE
 COMMENT Data kindly reviewed (02-NOV-1989) by Krumholz L. R.
 FEATURES
 source
 1. 7402
 /organism="Vibrio alginolyticus"
 /db_xref="taxon:663"
 /clone_1id="lambda dash"
 /clone="Va-3"
 promoter 143..148

APF-7

gene
CDS
/db_xref="taxon:666"
/chromosome="1"
/note="biotype: El Tor"
136..2307
/gene="VC2758"
136..2307
/gene="VC2758"
/note="similar to GB:M87049 SP:P21177 GB:M59368 GB:M74164
PID:145900; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="fatty oxidation complex, alpha subunit"
/protein_id="AAE95897.1"
/db_xref="GI:9657359"
/translation="MTQAKTLOVKOLANGIAELSPCAPSVNKLDTLHTLESIDKALD
ATAADSVKGLISSDKKAFIVGADTTEFLGPAKPEALDEMLPANKLPNKLELGA
PFTLSALKHNTLGCGCECYLATDFRIGDTSITGIPETKLGIMPGRGTVRLPRIGLA
DSAMETITQKACRAEELKVLGDLAVDSKLDISATITLTOAIEKLDWKRRQK
TSALTLSTLEAMMSFTMAKGVAVAGKHYPAPMTSVVTEEARLPKPAALDIKHA
FKLAKSTPAQALVGIFLNDQYIKGLAKOSAKAASODTQHAALVLAGINGGIAVQSA
LGAVPEAKMDIAPHSLEGMTFAKLLNKOLERGIDGFKMAGILASTIPSLHYAID
QADVLEAVENPKVKAVALSEVEGLVOTETILTSTPTINLAKSLKRPNCYLPFYP
HFPNPVHRPVEIINGEHTSEDTIRVAVYAAKKKSPILVNDCCPFYVNRPLPEYF
AGFSLMRDGAFTEDKVMKQFMPKGPATILDVAGIDTAHQAQVAAEGPPTMA
KSGREALDLYEAKKFGKNGSGFYQYTDKRGKPKKPSDVLALIAVPCGAPQSF
POTLIERMTIMINEVYLCIEGIIASQAEADMALVYGLGFPPFGVGRVYLDITIGIA
NVYAAEKYADIGALYQVPLKNNMAQGSFYSQAQVNSAL"
2333..3496
/gene="VC2759"
2333..3496
/gene="VC2759"
/note="similar to GB:M87049 SP:P21151 GB:M59368 GB:M74164
PID:145901; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="fatty oxidation complex, beta subunit"
/protein_id="AAE95898.1"
/db_xref="GI:9657360"
/translation="MNTVVIYDCLRTPMKSGKAGFRHQRHEDLSAHMKGILARNPO
VNKRETIYMGVCOOTLEGFNVARNALLAGLPEIGAVTVNRLOGSSMOLHDA
RAIMWGDAEICLYGVYEHMGRHVPTGVDHPDLKSNVAKAAMGCTATMCKINGI
SRQODEFAKSHARAHAAATLEGRFKEIILPTGSHADGTLPTLIDDEVIRPETTYAG
LAEIRVFPDANGTAVAGTSALSDASAMLVSEOKAKALGLTIRARIKAMVAGCD
PSIMGYCPVPAITHKALQRAGLTMDMDVVELNEAFSAOSLPCKADGLLEMDKVL
NGCAIALGHPGSGSTRISTTLINLMEAKDAKYGATMCGIGGIAITPERP"
NGCAIALGHPGSGSTRISTTLINLMEAKDAKYGATMCGIGGIAITPERP"
complement(3563..4471)
/gene="VC2760"
complement(3563..4471)
/gene="VC2760"
/note="similar to SP:P37598 GB:U00096 PID:1742732
PID:1742742 PID:1787949; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="transcriptional regulator, LysR family"
/protein_id="AAE95899.1"
/db_xref="GI:9657361"
/translation="MMPKSSLEMLDVAVALGSAEVLVLPKVPASISYVQRIESE
LGVVLRRLPRKVELTPAGELFMLEARQLRQMEIENQTKRAHGRKTLKTLIDVY
VKIDKMKPVEAVYQTFDAELQINMEVFGSWEALIAQGRADIVTGAFAVPGQDE
VRMGILDMAFVMSPNHPCEVREONLSGFTSQFLATCIDDTSSVLPKRREHYKQR
LLPNNYSATIESLKSGLGVGMPRHMAQPLLASQALVEKVLDPKELSHCCLVWRKD
NHLTIEMVAVYLGSSDQLYSDWIAQK"
4600..5829
/gene="VC2761"
4600..5829
/gene="VC2761"
/note="similar to GB:U00096 PID:1742733 PID:1742743
PID:1787950; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="multidrug resistance protein"
/protein_id="AAE95900.1"

gene
CDS
/db_xref="GI:9657362"
/translation="MANNDEHKMKVSPFOQYLVYLSALSMGFIATDMYLPAFKMETA
FATGEQIALSLTVFLVGMALQDLGLASDFGRNRLINLIVFTLASAGLASSST
VWOLLRLRIQAVICAPAVIWAQAVIKRYPSSVSOQLPATIMPLVSPALAPOLGV
VLADYFGMSIFLTLGMIYLVSTGQTNELIAPKETSIDQIKGLKSTYVIGNV
MMFASAAFPAYLGMPEIMAOGLQADIDLSFIPQIAPMGVGLKRAVQKYG
EVLKQIALFTLSAAMTFVASQMTLSSILAPFCLIAVANGAMYPIVVRALSSA
KOSPATAGLQNSLQISVSSLSALVMAASQAVPTGTAIAYLCLGGLMGVLSNRT
LREQFTTPDNRVVSDE"
complement(5891..7252)
/gene="VC2762"
complement(5891..7252)
/gene="VC2762"
/note="similar to SP:P17114 PID:290578 PID:43267
PID:290579 GB:U00096; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="UDP-N-acetylglucosamine pyrophosphorylase"
/protein_id="AAE95901.1"
/db_xref="GI:9657363"
/translation="MKFSYVILAAKGTBMHSMNPVLTHTLAKPMAKHVIDPCNNLG
AONIHLYVGHGSDMOQALANENVNVVLOAQIGTHAVDQASPHRODEKTLVYGD
VPLISDPTLESLEAQPTDGIALLTVLEDPTGQGRIVRRGPVVAIVQKDASEOK
LKEVNTGVIVANGRDLKRMKRLAGLNNNNAQGEYVLDVIAAARDEGRAVEAVPSHSI
EVEGVNDRQLALAELEFQARQAKKILBQGVMLRDPARDLREGTCCGSDVEIDVYI
IEGVNSIGNNVYVAGSILKDEIDDNVIRPVSVEGATVGENCTVGPFTLRPGAE
LRDVAHGVNFVEEMKRNARLGEGRKANRLTYLGAELIGVNVAGVITTCVNDGANKKT
VIDGDVFGSDCLVAPVTIIGNCATIGACTTTLKVNAAEGELVYTRAPERKIAGMORPA
KKK"
complement(7499..7921)
/gene="VC2763"
complement(7499..7921)
/gene="VC2763"
/note="similar to SP:P12988 PID:48340; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="ATP synthase F1, epsilon subunit"
/protein_id="AAE95902.1"
/db_xref="GI:9657364"
/translation="MAITPHLVNVSSEKLLFGLVETFOVTSSEGLGFIHGHTPL
TAKPQNRVIRYKLGHEHYVYSGVIEYVOPGATVADVAINGEDLDAAKAEAKR
AEERILQGDMDPQASLALAIQALRYIELTKRR"
complement(7940..9343)
/gene="VC2764"
complement(7940..9343)
/gene="VC2764"
/note="similar to GB:X16050 SP:P12986 PID:48339;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="ATP synthase F1, beta subunit"
/protein_id="AAE95903.1"
/db_xref="GI:9657365"
/translation="MARGKIVQITGAVDVEFPOSEYPSYDIALNVDSKERLVLEVO
QOLGGVIRIVMSSDGLRRGMTVONTGAPISVPVGTGLGRIMNVLSAIDERGKI
GAEVYSIRHPAPSYEEOSSATLELLETGVNIDICFPAGKRIKGLFGAGVGTVM
MELINNAIOLHSGISVAGVGERREGNDVYHEHGOEVNVNVEQELSKYAMVGOVM
EPGKNRVALVLTGCTAERKPRDGRDVLFLINDYRYTLAGTEVSALGMPSPVVGQ
PTLAEEGVLIQERTITTKKSTYSVQAVYVYPADDLTPSPATTPFAHIDAVVNRNIA
AMGIVPAIDPLDSTRLQDLVVGQEHYDVARQVATLQRIKELKDITAILGDELS
ADKQVAVARRKIERTFLQPYHVAEVFQDGVVYPLKETLRFKGLLAGDYDDIPROA
FYRCGTIDDAIENAKKL"
complement(9378..10244)
/gene="VC2765"
complement(9378..10244)
/gene="VC2765"
/note="similar to SP:P12990 PID:48338; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="ATP synthase F1, gamma subunit"

```

Db 4475 CCAGAAAATTCAGAAAAGGATGAAATGTAAGACACTTCAAAATAATGCTGTTCTCAAG 4534
Qy 473 aacgcgtctcttcacgccttattcagaacaataagtaagcgtatgacgaagttt 532
Db 4535 AAGCAATGAAAGACCTTCCTCCATACGATGCAAAACAAATGGCTAAGTCGATGCG 4594
Qy 533 ccaaaagcaacgattggttcaagcatcatttttaagtgatcgcgaagttaaaaaagttg 592
Db 4595 CTAAAGCAAACTAGAGTACCGTCACTCCCTACCTAGAAAGCGTGAAGCTAAACGTTGG 4654
Qy 593 gcaatgattgtgtccacagaatcgtgtcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 652
Db 4655 GTTACATCATCTTCGTCAGACAGCCGTCGTCGTGTGTGTGTGTGTGTGTGTGTGTGT 4714
Qy 653 aaacgtatataatgaatgaagaatggaagaatgaatgtttccgttcaattgatt 712
Db 4715 AAAAAGCCCTTACAGACAAACATGAGAGAGAAAGGTGCTGAGATTGACCTTGGCCG 4774
Qy 713 taatcggtcttaataatcattcaacttttccaaatcttgggaatttaaatttaaccgaag 772
Db 4775 TATGTTGGCTCAAAAGCAACAGCTTCTTTAAACATGTGTGGCCGAAAGTTGGCCGACAG 4834
Qy 773 attcaggtattgtgtatctccctctgtgtgacgaattatgtgttcaatctatga 832
Db 4835 TTTCTGGCCCTGGGTGATTAACCAAGCTTAGAAGACCTAATCGCTTGTGTGGCTAATGC 4894
Qy 833 ttgatgctataaaaaagggaagtagatgtgtgtatattagtttatacaaatattat 892
Db 4895 TCAAGAAATACATGAAAGGTGAGCTGACCGCTATATACGATTTCAACAAATTTGTGA 4954
Qy 893 aacagatgtcgcaaaaagcagattgtgaaaataattaccattaccagaattagataatg 952
Db 4955 ATACATATGTTTACGACACCAAGATGATCAATTTGCTACCTTTGCTTAATGGGACAGCG 5014
Qy 953 atgaattaggcgaaagaaacagtttggatataattacgaacctgtatgcyagaagat 1012
Db 5015 AAGAGAT---GCAGCGTGAGCACTCATGTGGACATCATATGAGCCTGAGCCAAACCTC 5071
Qy 1013 tatttgataattattgtgtcttatttgaatctcagtttatacaagcagcagttgaaa 1072
Db 5072 TACTGATATCGGTACTACTAGTGTGCTTACGATGAGTCTGAGGTATACAAAGGTGTGTGA 5131
Qy 1073 acctgtctctgagcaagcgcgtcgatgtgcgcatagaagcaagcaagataacgacg 1132
Db 5132 ACCTTGCTTGTAAACAGAGCGCTCGAATATTCGATGGAAGCTGCAACGGATTAACGCCA 5191
Qy 1133 gtaacttaataatgattacagttatgctatacaaaaagctcgtcaagaagattataca 1192
Db 5192 CCAACTGTATGACGATTTGGAACCTTGTACAAACAAAGCCGCTGAGCTGCGATCACAC 5251
Qy 1193 atgaattaaatgaattgttgcgcgtgtgcagcagcaattaa-----C 1234
Db 5252 AAGAAATATTCGGAATACGTTGGTGTGTGAGCTGCGGTTAAAGCTTAGTAAACGAATA 5311
Qy 1235 aataagagatcgtaattgtgcaactgtgaaataattgacaatatcatcggtgacgttata 1294
Db 5312 GTTTAAGAGATTAAAGATATGCTTACAGATGATGATGATGATGATGATGATGATGATG 5371
Qy 1295 cgttgaattccacagaatgacgtacacaaatataatgtccttaaatgt----- 1345
Db 5372 CGTAGAGTTCCACAGAGAAATGATACCTAGTATATGACGCTCTAAGCGTAAACGAGAC 5431
Qy 1346 tgaataaagtttagtactctgaagttcaacaacaataagttgtgtgtgtgtgtgtgtgt 1405
Db 5432 AAAAGAGCCTTATGTTCTTGAAGTTCAACAAACAGCTGAGCGGCGGTGATGTTGTTAT 5491
Qy 1406 cgcgaatgggatactcgtatgatttaaaacgcggttaagcgtaacaataagcaataacc 1465
Db 5492 CGTAAATGGTACGCTCTGATGTTGTTTACGCTGTGAGATTGAAGTGGTTAAACACTGGCC 5551
Qy 1466 aatttcgttcagltggaagcaaaaacatttggctcatcatgaacgtatttgggtgaacc 1525

```

```

Db 5552 AATTTCAGTACAGTGTGACTTAAACTCTAGTGTCTATCATGAAAGCTCTAGTGACGC 5611
Qy 1526 aatcgtatgcaagagtgaaatcgtgtcagaagaagattgttattcacgcgtgcacac 1585
Db 5612 GATTGATGAGCGTGTGTAAGTGGCCGGAAGAGTTTACTTATCTACACCGACGCCACC 5671
Qy 1586 aagttatgaagaacaactacactgaactgaacttttagaagaacggaattaaagtatacga 1645
Db 5672 AACTGCAAGAAACATCAAAAGAGATTCGACTTATGAAAGCGGCTTAAAGTAAATTCGA 5731
Qy 1646 cttaagtttccgttctgcgaagaagggttaagaagttattcctgtgtgtgtgtgtgtgt 1705
Db 5732 CCTAATTTGTCCATTTGCTAAGGCGGTAAATGCTTATTTGGTGTGTCAGGTGTAGG 5791
Qy 1706 taaacgcataatgattgaattaaatccgaatacgcgaatgacactgaactgaacttc 1765
Db 5792 TAAAGCCGTTAAATATATGAACTTATCAACAAATGCGCTACACAGACTCAGGTTATTC 5851
Qy 1766 tgccttcgaggggtagtgagcgtacgcgtgaaggaagtaacgacttcatatagatga 1825
Db 5852 CGTATTTGGCGGTGTGTGTGAGCGTACTCGTGAAGGTTACGATTTCTACAGAGATGCA 5911
Qy 1826 agacttcaagttataga-----taagtgcttctgtttagt 1864
Db 5912 GGAAGCGGCGCTGTAAAGCTTGAAGGCTGAAGATGCAAAAGTACGATGCTTACCG 5971
Qy 1865 tcaatgaacgagcgcaagcaggttaacgctttagtgatggaattacagcgttaactatg 1924
Db 5972 TCAGATGAACGAGACCCAGAGTACCGTCTACGTTGTGCACTACGTGTGTTAACAATGGC 6031
Qy 1925 ggaanaatttcgtatgaagtgatgtcttattcttcttctgttataa 1972
Db 6032 AGAAGCTTTCCTGAGAGAGGTGCTGACGTACGTACGTTCATTTGATTA 6079

RESULT 13
AE004342/c
LOCUS
DEFINITION
Vibrio cholerae chromosome I, section 250 of 251 of the complete
chromosome.
ACCESSION
AE004342 AE003852
VERSION
AE004342.1 GI:9657358
KEYWORDS
SOURCE
ORGANISM
Vibrio cholerae.
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
1 (bases 1 to 12928)
Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L.,
Dodson, R.J., Haft, D.H., Hickey, E.R., Peterson, J.D., Umayam, L.A.,
Gill, S.R., Nelson, K.E., Read, T.D., Teitelin, H., Richardson, D.,
Ermolaeva, M.D., Yamathayan, J., Bass, S., Qin, H., Dragol, I.,
Sellers, P., McDonald, L., Utterback, T., Fleischmann, R.D.,
Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R.,
Mekalanos, J.J., Venter, J.C. and Fraser, C.M.
Direct Submission
Submitted (14-JUN-2000) The Institute for Genomic Research, 9712
Medical Center Dr. Rockville, MD 20850, USA
FEATURES
source
1..12928
/organism="Vibrio cholerae"
/strain="NI6961"
/serotype="O1"

```


	gene	/protein_id="AAF95904.1"
	CDS	/db_xref="GI:9657366"
		/translation="MAGAKEIRTKIGSVKSTOKITTKAMENVAASKMRSDAMESRR YAQIRIKVIVGHANASLEVRHPYIESEBAKAVGYIITSDGLGGININIFKAITID MOMFKEGAOIELAIIISGARAFPNNSGAARAAVOYSIGSGSPSIEDLISGVNKKTY DKGELDLRLYLENFNVNTWPKRIDQLPLPKSDSEDMQRDHMDYITEPEPFLIED ALLRIETSOYGVENVNLACEQARNRVMKAADTNASNLIIDLVLNRNARAATTO ELSEVGSAAY"
		complement(10289..11860)
		/gene="VC2766"
		complement(10289..11860)
		/gene="VC2766"
		/note="similar to SP:P12985 P1D:48337; identified by sequence similarity; putative"
		/codon_start=1
		/transl_table=11
		/product="ATP synthase F1, alpha subunit"
		/protein_id="AAF95905.1"
Query Match	39.4%; Score 777; DB 1; Length 12928;	
Best Local Similarity	63.8%; Pred. No. 3.7e-142;	
Matches 1294; Conservative	0; Mismatches 665; Indels 68; Gaps 5;	
Oy	11 ttggtcagttcggttcggatcttctaigtacagtgtaagaacttatcacatggtcaa	70
Db	10615 TTTGTCTCAGTTCCTCTCGATCGATCTGGATGAAGGCACCAACGCTCACCTGACCTACCGTGAG	10556
Oy	71 aaagtcaatgataatttaaacaanaaccatactccgtatctctgtagcacaaagca	130
Db	10555 AAAGGACTGTAACATCATATAAGCAGAACAAGTAGACCAATCTCTGTTGATATAGGCT	10496
Oy	131 ttagtgtattctgcagtagagtttggttactaagaagaagtgacctagatcgtaattg	190
Db	10495 CTGGTGATCTTGGTGGCGGAACGTGGCTATGTGACAGACGTGAAGTGAACAGGTTCTC	10436
Oy	191 tcatttgataagcaacttttagagtagtctaaccaactatgctgatlttatgcgtgag	250
Db	10435 GACTTCGAAGCTGCTCTACTATCGATCTCTCGTCTATTAGCGTAATTGGCAGCGCAA	10376
Oy	251 ttaaccaatctgggaatatcaatgaatgaatgaatgaatgaatgaatgaatgaatga	310
Db	10375 ATCGCAAGACGGGGGCTTACACACATGATGATCGAAGCTCAGCTTAAGAACGCTGGTTGAC	10316
Oy	311 agcttcaagaagcaaaagtcgtygtgtaa-----gttaacactttaaat	352
Db	10315 GATTCCAAAGCAACCACCAACCTTGTAATGAGGTGCGTGGCAGTTTCTGCCACCACTAAC	10256
Oy	-333 ggaagacaaaatgycgaagggtctaaagagataagaaccaaactgcgagtgtaaaga	412
Db	10255 GGAGGTAAACATGCGCCCGCGCAAAAAGAGATACGTACTACAGATCGGAGTGTGAANAACA	10196
Oy	413 cacaanaaatattcaanaacgaatgaaatggttgcgtcgctcgaaatlgctaaagaag	472
Db	10195 CTCACAAAGATTACGAAACGATGGAATGAGTGCGCTTCAAAAAATGGGTGCTTACAGG	10136
Oy	473 aacgaatgctctctcaagccocttaattcaagaacaatacgtlaagtgtaacacg	532
Db	10135 ACGCATGAGATCATACAGCCCTTACGCAACAACATACGTAAATGATCGTCAATGTCG	10076
Oy	533 ccaaaagcaagatttggttacaaagcatccaattttagtgatcgcgaagtaaaaaag	592
Db	10075 CCAAGCAAGTCTAGAGATATCGTCAATCGGTATCTGAGAGAGCGTGAAACGAAAGCAAGTTGG	10016
Oy	593 gcataatggttggtgccacagatcggtgcttctgtggtggtcttaaacgtgaactgt	652
Db	10015 GTTACATATATTATTTCTACAGACCCCGGTTGTGTGGGTTTGAACATTTAACTTTGTTCA	9956
Oy	653 aaactgataatgaatgaagaatggaagaagaagaatgcttcogttoaattgagtt	712
Db	9955 AAAAAGCATCATCGGATTTGCAAACTTGGAAAAGAGAAAGGTGCGCAGATTGAGCTGGCGA	9896
Oy	713 taatggttctaatactatcaactttttccaactcttgygaattaataatttaacc	772

Dd	9895	TTTATGGCTGCAAAAGCGACCGCTTTTTCACACACAGCGGCCGAAATAGACGCTCAGG	9895b
Qy	773	atcagagatcgtgagatcactccctcgtctgtagcagctaaatctggttcaglucaattcattga	832
Dd	9835	TTTTGCGTTTGGGGACAGGCCAAGCTTAGAAGATCTGATTCGGTTGGGCGTCAAGC	9776
Qy	873	ctgatgcctataaaaaaggaggtagatgctgctattagttattacaatttat	892
Dd	9775	TGAAGAAATATAGATMAAGGATGTAGACCGCTCTATCTGGGTGTTCAACAGTTTGGA	9716
Qy	893	acacgatcgtccaaagccagatcttgaaaaattaatccattccattcaagaattagaatg	952
Dd	9715	ACACCATGTGTCAAAAACCAAAATGATCATTTGCTACCTTTGGCCAAATTCGACACCG	9656
Qy	953	atgaattaggcgaaagaaaacaagcttgtagatataattcagaaacctgatgcgaaat	1012
Dd	9655	AAAGATAT---GCACCGTATACATGTGGGACTACATCTAGACGCTGACCAACCAAGCTC	9599
Qy	1013	tattagataattatctggttcgttatttagaatctcaggttattacaagcagttgaa	1072
Dd	9598	TATTAGATGCGCTTTTACTCGCTTCATTGAGTCGCGAGTTTATCAAGTGTGTGAA	9539
Qy	1073	accttgcttcgagcaagccgcctcgaaatgctgcctatgaaagcagcaacaagtaaagcag	11322
Dd	9538	ACCTTACCGCTGTGACCAAGCACGCCGATGGTTCCGATGAAACCTCTACCGATPACGCGA	9479
Qy	1133	gtaaccttaataagagttacagttacgtatcatalaacaagaactgcgtacaagaattaca	11922
Dd	9478	GCAACCTGATTTGATGATTTTACAACTTGTGTACACAAACCCGTCAGGGGGGATTTAC	9419
Qy	1193	atgaattaaatgaatctgttcgcgtgtgcagcaattaa-----ca	1235
Dd	9418	AAAGATTGTGCGAANTGTGAGCGCGCGCTCGCTTTAACCTTAGTAAACAAACAG	9359
Qy	1236	aatgagagatcgtgtaagcgcaactgtgaaaaattgttcaaatcatcgtgtggttatgac	1295
Dd	9358	TTTTAGAGGATTTACGATGTGCTACAGGTACAGATGTACAAATCATGTGGTGGTACTGAC	9299
Qy	1296	gttgaattccccaacaagactgcagtaaccaaaagtalatgaatgccttaaatgttg-----	1347
Dd	9298	GTAGGTTCCACAGACAGACCAAGTACCAAGTGTACGATGCTGATAGTTGTGTGATTC	9239
Qy	1348	-aaacaggtttagtaacttgaaagttcaacaacaatlagtgtgtgtgtatgcgtatc	1406
Dd	9238	AAAGACGCTGTGTTCTGTGAAGTTCAACAGCACACTGCGCGGTGATTTGCGCCTATC	9179
Qy	1407	gcaatggatcatctgtgtatgattaaaaagcggtttagaagtgatacaaatatagaataacca	1466
Dd	9178	GTTATGGGTAGCTCGATGATGGTTTACGTCTGGGATACAGTGTCAAAACACTGCGCTCCA	9119
Qy	1467	attctgttcacgtgggaaagaaacaatctgggtgcgtatcatataagcttatcttggtgaaca	1526
Dd	9118	ATTTCAGTACCAGTAGTACTTAAACCCTTAGTCCGTATATATAACAGTCTTGCTGATGCG	9059
Qy	1527	atcgaatgacaaggttgaaatcgtgtgcaagaagaaatctgtctattaccogtgcgcaaca	1586
Dd	9058	ATTGACGAAGCTGGCGACATTTGGCGCAAGAAAGAGTACTCATTCACGCTCTCTCCA	8999
Qy	1587	agttatgaagaacaactcaacagtaactgaactttagaagaagggaaataaagtatatcgac	1646
Dd	8998	AGCTACGAAAGAACGTCTATGTGCACATCTTTCGAAACCGGCTTTAAAGTTATTCAC	8939
Qy	1647	ttagtctgccttggttgaaagggtttaaagtagtcttatctggtgtgtgcggtgtcggt	1706
Dd	8938	CTGATCTGTCCGTTTGGCAAAAGTGTGTAANTCGTCTTTTGGGTGTGCGGGTGTAGCT	8879
Qy	1707	aaaacgtcaaatatgatcgaatcaatccgttlaaacatgcgaattgagcaatcagttactct	1766
Dd	8878	AAAGACGTTAACAATGATGAACCTTATTCACAACAACATGCGCCTACAGCACTCAAGTTTGCA	8819
Qy	1767	gtcttctgggggttagtgtagcgctacgctglaaggtlaacagcttctatcatgatgatgaa	1826
Dd	8818	GTATTTCGTGGGTAGCTGTGACCTACTCTGTGAGGGGTAAACGACTTCTTACACGAATGACG	8759

```

/translation="MRNSTEISKLIERIAOEVEYNOSYNCSIIISVSDGIKINGL
SNVMGEMILPNNNEYATAINERDPTGAVVMPYIHI SEAKVCTCKIIEVDPGN
FLGRVYNAAGPFDIGKSDSIQNDGPPPVADAPGVIDRKSVPDPTOTGKVIDSMAPIC
KQGRLLIGDRQIGKCTALADIT1INQOSGICITVAIGQKSTIIINVKLENNAL
FNTIIVASASEAASLOYLAPYSGCAMEFENKEDSLIYDLSKHAAYRQISL
LRPRGRAPGDIIEYLSRLIERASRSMVEYOKITNKTIGKSTIAP1IETOS
GDSAFVFNYSITDGOIFLESNLFNSGIRPAPVPGISVSVGSAAGTIIKIKLSG
IRTAAYOELAAFSQSSDIDPTKRLNQCOKITELLKOKOYSP1IABALKEP
AENNFUDISIDKIKFEKELIYAHNNTYFDIMEIKNTGDPNIVIKDFIKLIEFK
KNQ"
```

gene
/gene="atpg"
6101..6973
/EC_number="3.6.1.34"
/note="BU007"
/codon_start=1
/transl_table=11
/product="ATP synthase gamma chain"
/protein_id="BAB12735.1"
/db_xref="GI:10038700"

gene
/translation="MTSTKEIKNKIVSVTNTKKTITKAMEBVAVSMKRTTEERMSGR
YSDIRKVIDHTQGNLEKHSYLEERTNRMGM1ISTDRLCGLMTNLFKQVLFK
IONFAKVIIPCDLIEFKLSVFKCSNIIAKATNGENFLEELINSVG1IIQY
OCKRIDKLEIAYNKNFHNKMSOYPTITOLLPSKKNODASNNMDDYIYEPESKLIIDT
LENYIESOVOSTIENIASHEAAMIMKKTATDNGSRIRIKELQIYKVKRQANTOE
LNEYSGASVNSID"
6997..8394
/gene="atpd"
6997..8394
/EC_number="3.6.1.34"
/note="BU008"
/codon_start=1
/transl_table=11
/product="ATP synthase beta chain"
/protein_id="BAB12736.1"
/db_xref="GI:10038701"

gene
/translation="MATCKIIQIGAVVDVEPNDSVKRIYNALEEVKKOKKILIEVO
QOLGSGVVRITAMSSNGLKRLIVTDLGHITKVPGEATLKR1IINVAETIDNKGAL
KNNSDKLEWEIHRSPSPYDQASCRLIETGIVIDLICPFSGGKVGLEGGAGVG
KYNMMLIRNIAVHESGYSVTGASCRILENDPZYHENDSOVLDRKSLVYGONNEP
PGRILRVAFGTITAEKPRDEKDVLEFDNIYRTYLAGTEVSALLGNRPSAVGYOPT
LAEBGLQERTSTKNGSITSVOAVPADLIDPSPATTFPAHLDSTVYTLSSROIASL
GIYPAIDPINSRQIDPYIVGDEHYETALGQSTLQKROELKIDITAILIAGDELAKD
KILSRARKIQEISFPFVAEFTGFPEKIVSLKDNIRAFKGIILKEGFDLPEOAFY
MVSIEEYIERAKKL"
8421..8837
/gene="atpc"
8421..8837
/EC_number="3.6.1.34"
/note="BU009"
/codon_start=1
/transl_table=11
/product="ATP synthase epsilon chain"
/protein_id="BAB12737.1"
/db_xref="GI:10038702"

gene
/translation="MNFYLDVYSLTKTIFSGFEVKIRVSGSEGLGIYPGHAQLSIL
KPGWYIHRKKDKKECIYISGILEOPEVSYSLADVAIHIDIDRSRIILTKKNAE
ESIKSNMKIKMDALLOISKEIRAKRLVLEWDRK"
complement(8911..11322)
/gene="gyrB"
complement(8911..11322)
/EC_number="5.99.1.3"
/note="BU010"
/codon_start=1
/transl_table=11
/product="DNA gyrase subunit B"

CDS
gene
Query Match 38.8%; Score 764.8; DB 2; Length 347550;
Best Local Similarity 64.2%; Pred. No. 7.7e-140;

```

Matches 1267; Conservative 0; Mismatches 657; Indels 48; Gaps 6;
QY 43 cacagtaagcaattatcacatggtcctaaagaactgattatgaacacaaacata 102
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5777 CACTAGAAAGCAATTAATCATGCTCAAAAATTAACGAAATTAATTAACAAAACAAATA 5836
QY 103 ccttcggtatctgtagcacaacaagaactagtgattatgtagaagattggttacct 162
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5837 TTCACCTATTTCTATAGCCGACAGGCTCTTACTTTTGTTCGCAAAAATTAATTTTCT 5896
QY 163 aagaagaagtgacttagatcgtatggtcattgtaacagacatttagatagctaa 222
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5897 TGATGATATTTCTATAGTAATAATCATTAATTTGAAAAAGAAATTAATTAATTAATCTCA 5956
QY 223 ccaatactatgctgattatgtagcgtgagttaaaccaactcggcaattacaatgagaat 282
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5957 TAATTTATTTTGTGATTAATGAGAAATTAATAAAGCAGGATTAATTAATTTGTTAT 6016
QY 283 taagaagtcattaaagagcatttggatagcttcaagaacaacagtcggtgaaattaa 342
    |||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6017 AAAAGTAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6076
QY 343 caacttaaatgagagacaaa---tgacaggtgctaaagagataaagaacaaatcgcg 399
    | ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6077 CTATTAAGTTAAAGAGAAATAGTGACGACGACAAAGAAATTAATAAACAATAATCGTG 6136
QY 400 agtgtaaaaagtaacaaaaaattactaaagagatggaattggtgctcctcgaatg 459
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6137 AGGTGTACCAATACATAAATAATCATTAAGCATGAAATGCTGCACTGCTTAATAAGT 6196
QY 460 cytaaacgcaagacgacgctgcttcacgacctatcagaagaacaatagctagctg 519
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6197 AGAANAACCTGACAGAAAGAAATGCGATCAGCTGACCATATTTCTGATATTATTAGAAAAGTT 6256
QY 520 attagccaagtttccaaagcaagatggttacaagacatcattttagtggatcgcgaa 579
    ||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6257 ATTGATCATGTGACACAGGAATTTAGAAATTAACATAGTTACTTGGAGAGAAAGAAA 6316
QY 580 gtaaaaaagtgagcagatgattgtgtgtccacagatcgtgcttgggtggtcctaagc 639
    || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6317 ACTAACCGTATTCGCATGATTAATTAATCTACTGATGAGAGGATTAATGCTGCTGTTAAAT 6376
QY 640 gtagaacttgtaaaactatataaagaatgaagaagaatggaagaagaatgcttc 699
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6377 ACTAATCTTTTAAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6436
QY 700 gttcaattgagtttaactcgttcctaaactcactcaactttccaaactcttggaaataa 759
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6437 TGTGATTTAATTAATTTGTTTAAAGATTATCTGATTAATTAATTTGTTGGAAGTAAT 6496
QY 760 attttaaccaagattcaggtatggtgatactccctcgttgagcagttatggttca 819
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6497 ATCTTGCAGAAAGCACTAAGTAGAGAAATCCCAAAATTAAGAGAAATTAATTAATTTCT 6556
QY 820 gtaactcattgatgtagcttataaagaagggaagtagatggtgattatgattat 879
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6557 GTTGGATTTATCTCCAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTTATGCTTAT 6616
QY 880 aacaaatttataacacgagtgcgcaaaagcaglatctggaataataatcattaca 939
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6617 AACCAATTTCTATTAATAATGTCGCAATATCTACAAATTAACCAATTAATTAATTAATTTTCT 6676
QY 940 gaattagataatgataatgagcgaagaagaagaatggtgattatattcagaact 999
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6677 AAA-----AAAATGATCAAGATGCAACCAATTAATTAATTTGATTAATTTGATGACCA 6730
QY 1000 gatgcaagattatataatattatggtgctgattatggaactcagaattatca 1059
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6731 GATCTTAACCTATTTTAAACACTTTTAATCGATTAATTAATTAATTAATTAATTAATTAATTA 6790
QY 1060 gcaagcagttgaaacactctcctcgtgacagcgcgcgagatggtgcacatgaagaagca 1119
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6791 AGCATTTTGAAGAAATTAATGACAGTGAACAGCTGCTGATGATGATGATGATGATGATGATGATG 6850
```



```

Db 231 CTGGATATTAACACACCTTACCTGGAAGACCCGAGCTTAAACGCTGGGCTACCTGGTG 290
Qy 604 ggtccacagatcgtggtcttctgtgtggtcctaaacgtgaactgtttaaactglat 663
Db 291 GTTGTGACACCGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 350
Qy 664 aatgaatgaagaatgaagaagaagaagaatgttccgttcaattgaattgaattcgttct 723
Db 351 GCGGAATTAACAACTGAGACCGACCAAGGCGTTCATGCGACCTGCAATGATCGGCTTG 410
Qy 724 aaatcatcaactcttcccaatccttgggaattaaattttaaacaagaattcaagtatt 783
Db 411 AAGGCGGTCTGCTTCTTCAACCTCCGCGGCGCAATGTTGTTGCCAGTCCACCGCATG 470
Qy 784 ggtgatactccctcgttctgaagcaatgaattgttcagttcaattcttatagtatgtat 843
Db 471 GGGGATTAACCTTCCCTGTCGCAACGATGCGGTAAAGTGATGTTGACAGGCTTAC 530
Qy 844 aaaaaaggggaagtagatgtgtgtattagtattataaacaatttataacagatgtcg 903
Db 531 GACGAAGGCGCTGTGACAAAGCTTACATGTCAGCAACAAATTTATTAACACCATGCT 590
Qy 904 caaaagccagatttggaaaaaatattcattacagaaattagataatgaattagagc 963
Db 591 CAGGTTCCGACATGACCGACGCTGCTGCCGTTACCGGCAATCAGATGATGATGCT---- 646
Qy 964 gaaagaaaaaagatttggatataatttaacgaactgtacgtacgaagaattatagataat 1023
Db 647 --GAAACATTAATCCGAGGATTAACCTGTACGAACCGCATCCGAAGCGTGTGATAC 704
Qy 1024 ttattgttcgtatttagaatcgaagtttcaagcgcagttgaaacttgccttct 1083
Db 705 CTGCTGCGTGTATGTCGAATCTCAGGTTTATCAGGGCGGTGTAACCTGGCCACG 764
Qy 1084 gagcaagccgcgcgaatgttcgcataaagcagcaagataacgcaagtaacttaatt 1143
Db 765 GAGCAGGCGCGCGGTATGTTGCGATGGAAGCCGCGACCGACATGCGCGCATGATT 824
Qy 1144 aatgagttacagttagttcctatacaaaagctgcacaagtaattatacaatgaat 1203
Db 825 AAGAGACTCAGTGTGTATACAAAGCTGTCAGGCGACGATTAACGAGAACTCAC 884
Qy 1204 gaaattgttgcgcgtgacgaacttaacaa-----atagaggaatcgttaatg 1254
Db 885 GAGATCGTCTCGGGGCGCGCGGTTTAAACAGGTTATTCGTAGAGGATTTAAGATG 944
Qy 1255 caactggaaaaatttgaacaatcgcgtgtggttattgaagttgaattccacaagatg 1314
Db 945 CTACTGGAAGAATGTGTCCAGGTAATCGGCGCGGTAGTACGTCGAATTCCTCAAGATG 1004
Qy 1315 cagtaaccaaaatataatgactt-----aaatgttgaacaggtttagtactg 1365
Db 1005 CCGTACCGCGCGGTATGATGCTTGAAGTCAAAATGTAATGAGCGCTGTGCTG 1064
Qy 1366 aagttcaacaacaattaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1425
Db 1065 AAGTTCAGACGAGCTCGCGGCGGTATGTAATGCAATGCAATGCGGTTCTCGGACG 1124
Qy 1426 gattaaaaagcgtttaaagcgtlaacaacaataacgaatttctgttcagttgggaa 1485
Db 1125 GTCTGCGTCCGCGTCTGATGTAATAAAGACCTGGAACCCGATGGAAGTCCCGTAAAGTA 1184
Qy 1486 cgaaaaacatttgggtcgtatcatgaaactgtgtgaacaaactgtagaagaatgaa 1545
Db 1185 AAGCGACTCTGGCGGTATCATGAAGCTGATGGAACCGGTGACATGAAGAAGCGAGA 1244
Qy 1546 tgggtgcagaagaagaattgtgtatctaacctgtgcacacaaagtataagaacaacta 1605
Db 1245 TCGGTGAAGAAGAGCGTGGGATTCACCGCGACACCTTCCTACGAAGAAGCTGTCAA 1304
Qy 1606 aacagtaactttagaagaacgggaattaaagttatcgaacttagtttccgttgcga 1665

```

Search completed: October 28, 2001, 16:35:31
Job time: 5810 sec

```

Db 1305 ACTCTAGGAAGTCTGTGAACCCGCTATCAAAAGTATTCGACCTGATGTCGCTGCTA 1364
Qy 1666 aaggggtaaaagtagttattcogtltgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1725
Db 1365 AAGGCGGTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1424
Qy 1726 aattatcgttaacatgcaatttgcactgaactgaactgaactgttctgttctgttgcgtgt 1785
Db 1425 ACCTATTCTGTATACATCCGATCGAGACCTCGGTTACTGTGTGTGTGTGTGTGTGTGT 1484
Qy 1786 aagcagcgttgaagtaacgacttctatcatgaaatgaagaacttaacgtattagata 1845
Db 1485 AACGTACGTGTAAGGTAAACGACTTCTACCAAGAAATGACCGACTCAACGTATACGACA 1544
Qy 1846 aagtgctctgttattgtgtcaaatgaaagcagccacagtaacggtttagtgtgcat 1905
Db 1545 AAGTATCCCTGCTGTATGAGCGAGATGAAGAGACCGCGGAAACCGTCTGCGGCTTGC 1604
Qy 1906 taacagcttaactatggtcggaataattcgtgtatgaagaagtcgtgtatcttattctcg 1965
Db 1605 TGACCGGTCTGACCATGCTGTGAATAATCCGTGACGAAGTGTGATGCTGTGCTGTG 1664
Qy 1966 ttgataa 1972
Db 1665 TTGACAA 1671

```



xx Claim 1, Pages 67-68; 322pp; English.

ps The family Pasteurellaceae encompasses several pathogens that infect a
cc wide variety of animals. The present invention relates to virulence genes
cc from Pasteurellaceae. The present sequence is one such virulence gene.
cc The present sequence may be mutated in order to produce an inactive gene.
cc The inactive virulence gene may in turn be used to produce a vaccine,
cc which is useful for treating bacterial infections such as septiciemias,
cc bronchopneumonias, rhinitis and wound infections.

xx Sequence 1972 BP; 652 A; 312 C; 438 G; 570 T; 0 other;

Query Match 100.0%; Score 1972; DB 21; Length 1972;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1972; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agcggcattgtgctcagtcggtcgcgttcggtatctatgtagcacagtaagcaattatc 60
DB 1 agcggcattgtgctcagtcggtcgcgttcggtatctatgtagcacagtaagcaattatc 60
QY 61 acatggtcaaaaagtaactgtaattatgtaaaacaaactactctccgttatctgagc 120
DB 61 acatggtcaaaaagtaactgtaattatgtaaaacaaactactctccgttatctgagc 120
QY 121 acacaagcattgtgtatttgcagtagaggttggttacttgaagaagtggaattaga 180
DB 121 acacaagcattgtgtatttgcagtagaggttggttacttgaagaagtggaattaga 180
QY 181 tcgatgtgtcatttgaatcagacacttttagagtagtactaacataactatgctgatt 240
DB 181 tcgatgtgtcatttgaatcagacacttttagagtagtactaacataactatgctgatt 240
QY 241 tatgctgtgattacccaactcgtgcaattacatgtagaatttaaaagtcataaaagg 300
DB 241 tatgctgtgattacccaactcgtgcaattacatgtagaatttaaaagtcataaaagg 300
QY 301 catttgtagcttcaaacagacagtcgtgtagttaaacttaaaatggaagagac 360
DB 301 catttgtagcttcaaacagacagtcgtgtagttaaacttaaaatggaagagac 360
QY 361 aaaaatgacagtgcttaaaagataagaacaaatcgcgagtgtaaaagtacacaaaa 420
DB 361 aaaaatgacagtgcttaaaagataagaacaaatcgcgagtgtaaaagtacacaaaa 420
QY 421 attactaaagcgtatggaattggtgtgctgcgtcgcgaataatgctaaacgcgaagacgac 480
DB 421 attactaaagcgtatggaattggtgtgctgcgtcgcgaataatgctaaacgcgaagacgac 480
QY 481 tctcttcacgccttattcgaacaataacacgtaacgtaacacgtaacacgtaacacgta 540
DB 481 tctcttcacgccttattcgaacaataacacgtaacgtaacacgtaacacgtaacacgta 540
QY 541 aagaatggttaacagcatcattttagtggatcgcggaagttaaaaaagtggtgcatgatt 600
DB 541 aagaatggttaacagcatcattttagtggatcgcggaagttaaaaaagtggtgcatgatt 600
QY 601 gttggtccacagatcggtgcttcttggtggtggttaaaacgtaaaacttcttaaaactta 660
DB 601 gttggtccacagatcggtgcttcttggtggtggttaaaacgtaaaacttcttaaaactta 660
QY 661 ttaaatgtaaaagaaatggaagaaagaaatggttccggttcaatttgaagttaactggt 720
DB 661 ttaaatgtaaaagaaatggaagaaagaaatggttccggttcaatttgaagttaactggt 720
QY 721 tctaaatctatcaacttttccaatcttgggaattaaatttaaaccaagatcaggt 780
DB 721 tctaaatctatcaacttttccaatcttgggaattaaatttaaaccaagatcaggt 780
QY 781 attggtgatactccctcgttgtagagtagtaattggttcaagcaattcattgattgattct 840
DB 781 attggtgatactccctcgttgtagagtagtaattggttcaagcaattcattgattgattct 840

QY 841 tataaaaagggaagtagatgttgtgtattttagtttaacaaatttaaacagatg 900
DB 841 tataaaaagggaagtagatgttgtgtattttagtttaacaaatttaaacagatg 900
QY 901 tcgcaaaagccgatttggaaaataatccattacccaagaaatagaataatgataat 960
DB 901 tcgcaaaagccgatttggaaaataatccattacccaagaaatagaataatgataat 960
QY 961 ggcgaagaagaacaaattggtgattatattacgaactgtagcgaagtattattagatt 1020
DB 961 ggcgaagaagaacaaattggtgattatattacgaactgtagcgaagtattattagatt 1020
QY 1021 aatttattggttcgttatttagaattcaggtttacaaagcaggttgaaaacttgcct 1080
DB 1021 aatttattggttcgttatttagaattcaggtttacaaagcaggttgaaaacttgcct 1080
QY 1081 tctgagcaagccgctcgaatggtcgcgcatbaaaagcagcaacagataacgtaactta 1140
DB 1081 tctgagcaagccgctcgaatggtcgcgcatbaaaagcagcaacagataacgtaactta 1140
QY 1141 attaatgattacagttagttactataaacaagctcgtcaagcaagttatacaattat 1200
DB 1141 attaatgattacagttagttactataaacaagctcgtcaagcaagttatacaattat 1200
QY 1201 aatgaaattggtccggtgcagcagcaatttaacaataagagatcgttaattgcaactg 1260
DB 1201 aatgaaattggtccggtgcagcagcaatttaacaataagagatcgttaattgcaactg 1260
QY 1261 gaaaatgtgtacaaatcatcagtggtcgttattgaacgttgaattccccaagatgagttac 1320
DB 1261 gaaaatgtgtacaaatcatcagtggtcgttattgaacgttgaattccccaagatgagttac 1320
QY 1321 caaaagataatgaccccttaaatgttgaacaggtttagtacttgaagttcaacaacaaat 1380
DB 1321 caaaagataatgaccccttaaatgttgaacaggtttagtacttgaagttcaacaacaaat 1380
QY 1381 taagtgtgtgtgtagttcgtctgtatcgcgaatggatcgtatcgtatgtaaaacggtgt 1440
DB 1381 taagtgtgtgtgtagttcgtctgtatcgcgaatggatcgtatcgtatgtaaaacggtgt 1440
QY 1441 taagcgttaacaatacgaataaaccaaattctgtccagtggtggaagcaaaaacatggtgtc 1500
DB 1441 taagcgttaacaatacgaataaaccaaattctgtccagtggtggaagcaaaaacatggtgtc 1500
QY 1501 gtaacatgaacgtatgtgtgtgaacccaatcgatgagcaagtggaattggtgtgcgaagaaga 1560
DB 1501 gtaacatgaacgtatgtgtgtgaacccaatcgatgagcaagtggaattggtgtgcgaagaaga 1560
QY 1561 attggttattcacccgtgcgcacccaagttaagaagaacaaatcaacagttactaactt 1620
DB 1561 attggttattcacccgtgcgcacccaagttaagaagaacaaatcaacagttactaactt 1620
QY 1621 tagaagcgggaattaaagttatcgaactagtttgcgttgcgaaggggttaagtag 1680
DB 1621 tagaagcgggaattaaagttatcgaactagtttgcgttgcgaaggggttaagtag 1680
QY 1681 gttattcgtgtgtgcgggtgtcgtgtaaaacggtcaatatatggaatttaacccgttaaca 1740
DB 1681 gttattcgtgtgtgcgggtgtcgtgtaaaacggtcaatatatggaatttaacccgttaaca 1740
QY 1741 tcgcaattgaaacactcaggttactctgtccttgcggtgtggtgtagcgtgcgctgaag 1800
DB 1741 tcgcaattgaaacactcaggttactctgtccttgcggtgtggtgtagcgtgcgctgaag 1800
QY 1801 gtaacgactctatcatatgataatgaagaacttaacgtaattagataaagtgtctctgttt 1860
DB 1801 gtaacgactctatcatatgataatgaagaacttaacgtaattagataaagtgtctctgttt 1860
QY 1861 atgtgtcaaatgaaagcagccacgaagttaacggtttagcgtgtgtgcatataacaggttaacta 1920
DB 1861 atgtgtcaaatgaaagcagccacgaagttaacggtttagcgtgtgtgcatataacaggttaacta 1920

OY	121	tcatttgatcagcactcctttagagtaagtcacacatacctaagctgattttagtcgtgag	250
Db	234563	GCCTTTGAATCTGGAATTTTGTGAGCTTTGTCCGTACCCACATCCGGAAAGTTTGGAAAGCC	234504
OY	251	ttaacccaactctggtccaattacaatgtaagaaatbaagagcatctaaagagcatcttgagat	310
Db	234503	GTCAAATGCTTACAGTGCANHTCCGACGAGAGCGAGAAACCGTTGAAAGCAGCCATGAAA	234444
OY	311	agcttcaagcaaacagctgctggtgtaagtaacatcttaagtggagagaca---aaatg	366
Db	234443	TCCCTTCAAAATCTTCTTACGCCCTATCAGGCATTAAGCGTTGAAGTGAAGAGTGCTGAATG	234384
OY	367	gcaggtgctaaagataagaaccaacaatcgaggtgtataaagtacacaaaaatctact	426
Db	234383	GCAATAGGAAAAAGAAATTCACCCAAATCCGCAAGTGTTCAGATATCCCAAAAGATCACT	234324
OY	427	aaagcgatggaatgagttgtgcctgcctgaaatgqgttaaaaacggaagaagcatgctcttc	486
Db	234323	AAAGCGATGCGATGATGGTGTCAACTCTTAATAATGGGAAGCTCAGGAACGAGTGCCTTGG	234264
OY	487	tcaagcccttattagaacaacatacgtlaacgtgtatgacacgcttttccaagaacagat	546
Db	234263	GGGGTCCGTATGCCGAAAAAGTCCGATATGGTATGAGCATCTGGCGCAACCAATAC	234204
OY	547	ggttaacaagcatccactttt---agtgatcgcgaagtlaaaaaagtgggcatgattgt	603
Db	234203	GATCATGGTATTCGTTACTGGAATCTCATCGGGAATTCAGACGTGCGCTTTATTTTGG	234144
OY	604	gtgcaccaagatcgtggtcttlttggtggtggttlaaacgttgaactgtttaaactgatta	663
Db	234143	ATTACGTCGTGAATAGGGTTGTGTGGGGTTTGAACCCCAACGTCGTAAGAAAGTTTGG	234084
OY	664	aatgaatgaagaagatggaagaanaagatgcttcgtctcaattgagtttaactcggtctc	723
Db	234083	GCACAGTGTCAAGATTCGGAATCAACAGATTTTAAAGAGAAATCGATCTTTGGCAGT	234024
OY	724	aaatctatcaactcttttccaaactcttgggaattaaatlttlaaocccaagatlcaggtat	783
Db	234023	AAAGCTGTGATGGCGTGTCAAGCATTTGTCTGAAATGTGGTTCGCACTGGCCGTAATTTG	233964
OY	784	ggtgatctccctctgttgagcagttaatggttctcagttcaactatgattgattgcttat	843
Db	233963	GGCAGTATCCCCAAAAATGGAATTCCTCTCGGACCTTTGACAGAACTCTTCCAAACGGTAT	233904
OY	844	aaaaaaggggagtagatgattgtgtattctttagttataaacaattatlaacacgattcg	903
Db	233903	GAGAAACATGAAATTTGACAGAAATCCATCTGTGTATTCGGGTTTGTTCATATACCATGCT	233844
OY	904	caaaagccagtatgtgaanaaaatlaattccatlaacagaattlagaatagtatgaatgagc	963
Db	233843	CAAAACACGGAATAGGAAGTATTCCTCTCATCGTGAGAAACGATGGATGGCATTCAGCT	233784
OY	964	gaaagaaaaacagt---ttggattatattttagaactcgtatgcggaagatattatagat	1020
Db	233783	CCCCAATCACCGTTCACGTGGGAATACCGCTACCAACCGATCGACTTGCAGTGTGGAA	233724
OY	1021	aatttatggtcttattttagaattccagtttattcaagcagcagtttgaanaacctgtc	1080
Db	233723	TATCTGTTGCCCTCTATTAGATCTCTGTGTTATATAGGCGTTTGAACCGCAATATGGCA	233664
OY	1081	tctagcaaacgcctcgtgaatggttcgcacatgaagcagcaacagataacgcaggtlaactta	1140
Db	233663	TCCCAACAGCGACGCCGCGATGGTGTGCATGAACCTGCACACAGCAATGCAGCAATGCC	233604
OY	1141	attaatgattcaagcttagctctataaaaaaagctcgttcaagcaagattataaatgaatta	1200
Db	233603	ATCAAAAGGTTGCGCTTTGTATATAAACAATCGGTCACATCGCATTCACAGGAATTG	233544
OY	1201	aatgaatgtgttcgcgtgtcagcagcaattaa-----caatag	1240
Db	233543	TCAGAAATTTGACAGATGGCGCGCGCTGTCTGATGGCCGCTGGAATCTCTGACGGAATTA	233484
OY	1241	agatcggtaatggtccaactgtgaanaaattgttacaacatcatcgtgtcggtatttgaactga	1300

Db 233483	GGATACGATTAATGAGCCCAAGGCCAAAATCGTACAAATTAATGCTGCGGTTTGTACCTGGA	233424
Qy 1301	attccacaagaatgcagtaaccacaagaatataatgtaaccttaa---tgltgaacaaggttc	1357
Db 233423	ATTTCACAGCGACATGATGATTCGGCCGTTTACGACGCTTTGAATAATGAGCAAAAACGGTCT	233364
Qy 1358	agtaactgaagttacaacaacaataagtggtggtggtgtaagtcgctgtaatgcgaatggatc	1417
Db 23363	GACTTTGGAAATCCCAACAGCTTTTGGGCGACGGGGGTGAGTCCGTCGCCATTCGATGGGCGAC	233304
Qy 1418	atctgataagataaacaagcggtttaaagcgtaaacaataagcaataaaccaaattctcttc	1477
Db 233303	CTCGACGCTTTGGAACCGGCGCATAGCTGTGAGCAATACCGGTGCGCCATTAATCTGTGCC	233244
Qy 1478	agtgggaacggaacaacatagtggtcgtatcatgaagataatgttggtgaaccaatcgatagca	1537
Db 233243	GGTAGTAAAGGTACGTGTTGGAGCCCATTTGCGATGTTTGGAAACGCTTTGACGAAAGC	233184
Qy 1538	aggltgaatcgtgycagaagaagatlygltcattcaccgtygcgcaccaagaatgaaga	1597
Db 233183	AGGTCCGATTCGATACCGACAAGACCGCTGCCATTCACCAAGCTGCCCTTAAGTTTGACGA	233124
Qy 1558	acaactcaacgtaactgaacctttaagaacaagggaaatlaaagttatcgacttaagttgtcc	1657
Db 233123	ACTGCTTCCACCAACCGCATTTGCTCGAAACGGGGCATTTAAGTGAATTGACTTGCTGTGTC	233064
Qy 1658	gtttcggaagaagggttaaagtaagttatttcggtgtygcgagtygtcgtaaacgtlcaa	1717
Db 233063	GTTTCCCAAAGCCGTTAAAGTAGTCTGTTTCGGCGGTGCGGGTGTGGGTAAACCGTGGAA	233004
Qy 1718	tatgatagaatlaatccgtaacatcgcgaatltgaagcacttcagttactcgtcttcgcgag	1777
Db 233003	CATGATGGAATTTGATTCACAACACATTCGCCAAAGCGCACAGCGGCTTGCCTGTTTCGAGG	232944
Qy 1778	ggtaggtgaagcgtacgcgtgtgaaggttaacgactctcatatgaatgaagaactctaact	1837
Db 232943	CGTGGGTAGCGGTACCCCGCAAGGTAAAGATTCTTACACAGAGATGAAAGTTTCCAACGT	232884
Qy 1838	attgaataaagtgtctctgtgttlaatgttcaaatgaacgagccacagtaacgatttaag	1897
Db 232883	ATTGGATTAAGGTGGCATGTGTTACGGTCAAGATTAAGAAACATTCGCGGCAACCGTTTGGC	232824
Qy 1898	tgtgcatlaaacaggcttaactatagcgcgaaaaaatctcgtaatga 1943	
Db 232823	CGTCCCATTTGACCGGTTTGACCAATGCGGGAATACTTCCGTGACGAA 232778	
RESULT 4		
AAA81489/c		
ID	AAA81489 standard; DNA; 837096 bp.	
XX	AAA81489;	
AC		
XX		
DT	04-DEC-2000 (first entry)	
XX		
DE	N. meningitidis partial DNA sequence gnm_37 SEQ ID NO:37.	
XX		
KM	Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;	
KW	antigen; vaccine; diagnosis; infection; antibacterial; identification;	
KX	Meningococcus B; MenB; ds.	
XX		
OS	Neisseria meningitidis.	
XX		
PN	WO200022430-A2.	
XX		
PD	20-APR-2000.	
XX		
PF	08-OCT-1999; 99WO-US23573.	
XX		
PR	09-OCT-1998; 98US-0103794.	
XX		
PR	30-APR-1999; 99US-0132068.	
XX		


```

OY 847 aagagggaagatggtgtgtattagttataacaattattataacagatgctgcga 906
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1225 gaagaagaattggaattataataataataataataataataataataataa 1284
OY 907 aagccagattggaataataataataataataataataataataataataataa 966
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1285 aagccataatcagaataataataataataataataataataataataataa 1344
OY 967 agaaacaagttggaattataataataataataataataataataataataa 1026
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1345 -----cattgtctctatgaattgagccagataagaatactatacttaagtgaac 1398
OY 1027 ttggtcttattgaattcagttatcagttatcagagcagttgaaacctgtcttag 1086
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1399 ttgctcattatggtgaggtttagttatcaggaataatttagaagcaaaagttgag 1458
OY 1087 caagccgtcgaatggtcgcataagcaagcaagcaagcaagcaagcaagcaag 1146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1459 catgcaacacgtatgactgctgagtaaaatgcaactgataatgcaactgataat 1518
OY 1147 gattacagttatgctataacaagaactgctgaagcaagtaatacaaatgaatga 1206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1519 gactatacattgaatataacagagagagagagagagagagagagagagagag 1578
OY 1207 attgttcggtgacagcagcaatttaa-----caataagagatcggtatgcaact 1259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1579 attgttgggtgttcgcagcagcgttgaataataattaaagaggaataatgcaatggaatt 1638
OY 1260 ggaataatgtcaataatcgcgtgctgattgagcgttgaattccacaagatgcagta 1319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1639 ggcgtgtgaacccaagttatgctgctgaattgattgctgattgaaataacgaagt 1698
OY 1320 ccaaaagtatagtatgcttaaatgttga-----aacaggtta 1358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1699 cctaaataataacagcctgtgttattgattgctgaagagagagagagagagagagag 1758
OY 1359 gfactggaattcaacaacaattaggtgtgtgattgctgctgattgcaatgagatga 1418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1759 acattagaaggtgtgctgataatgaatgagagcgtgtgttcacaaattgctgattgta 1818
OY 1419 tctgattgattaaacgcggtttaaagcgttaacaataacgaataacccaattctgtcca 1478
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1819 actgattgtgtcacaagagcattgataaagaataacgaagcaaggaattgtagtaact 1878
OY 1479 gtgggaacgaacaattggtgtctatcattgaacgtattggtggaacacgcgtgacaa 1538
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1879 gtgtgtgacgaacaactagtcgttatttaattgactagtgaaacaattgaccttaa 1938
OY 1539 gctggaatcgtgacgaagagagatg---gtctattcacgcgtgcacccaagtattga 1595
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1939 gaagaagaattgagattcgttcgcgcgacactcattcaccgtaacgacacacattcgat 1998
OY 1596 gaacacatcaacagttacgttaacttttagaagaacggaataaagattacgtactagttgt 1655
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1999 gaacttcaacagagttcaaatlttagaacaaggtattaaagtagtattactacagga 2058
OY 1656 ccgttgcgaaggggtgaagtagttatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1715
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2059 ccttatttaaaggtgtgaataacgtattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2118
OY 1716 aatatgatgaattacccgtaacatcgcaattgagcaatgagctactctctgtgtg 1775
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2119 ttaataccaagaattataataacaacacgtcgaagcaagcaggtgtgtattcttctatcgcc 2178
OY 1776 ggggtgagtgagcgtgacggtgtaaggttaacgaacttcatcattgaagtaaaacttaac 1835
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2179 ggtgtgagtgagcgtgacgtgtgaaggttaacgatttactactcgaatggtgtacgtgtgt 2238
OY 1836 gttattgataaagttctctgttattgattgaatgaagagacacacaggttaacggttta 1895
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2239 gtaattagaanaaacgcagcgtgtatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2298

```

```

OY 1896 cgtgtgacataacagcttaactatgctgcgaataatccgtgattgaaggtctgtatgc 1955
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2299 cgtgtgacataatcgtgtgttaacaaatgctgataattcttcgtgacgaanaagttcaagac 2358
OY 1956 ttattcct 1963
    ||| |||
DB 2359 gttattatt 2366

RESULT 7
AA13153
ID AA13153 standard; DNA; 7750 BP.
XX
XX
AC AA13153;
XX
XX
DT 19-MAR-1999 (first entry)
XX
DE Enterococcus faecalis genome contig SEQ ID NO:216.
XX
XX
KW Enterococcus faecalis; contig; detection; Enterococcal infection;
KW vaccine; attenuation; computer readable medium; ds.
XX
OS Enterococcus faecalis.
XX
PN MO9850555-A2.
XX
PD 12-NOV-1998.
XX
PF 04-MAY-1998; 98WO-0508985.
XX
PR 14-NOV-1997; 97US-0066009.
PR 06-MAY-1997; 97US-0044031.
PR 16-MAY-1997; 97US-0046655.
XX
PA (HDMA-) HUMAN GENOME SCI INC.
XX
PI Barash SC, Dillon PJ, Kunsch CA;
XX
XX
DR WPI; 1999-045171/04.
XX
XX
PT New isolated Enterococcus faecalis polynucleotides and polypeptides
PT - used to develop products for the detection of Enterococcus and for
PT use in vaccines for prevention or attenuation of Enterococcus
PT infection.
XX
PS Claim 1; Page 1098-1102; 2084pp; English.
XX
XX
CC A computer readable medium has been developed which has recorded on it
CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
CC AA12938 to AA13191 represent these nucleotide sequences which are
CC primary nucleotide sequences, also known as contigs. The computer-based
CC system can identify fragments of the Enterococcus faecalis genome with
CC commercial importance. The products can be used to detect the presence
CC of Enterococcus faecalis in samples. They can also be used for
CC diagnosing Enterococcal infection in an animal and monitoring
CC progression of disease, and for identifying agents which can be used to
CC modulate the growth or pathogenicity of Enterococcus faecalis, or
CC another related organism, in vivo or in vitro. In particular the
CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences
CC can be used in vaccines to prevent or attenuate an Enterococcal
CC infection.
XX
XX
SQ Sequence 7750 BP; 2520 A; 1373 C; 1631 G; 2216 T; 10 other;

```

Query Match 17.0%; Score 335.2; DB 20; Length 7750;
 Best Local Similarity 55.3%; Pred. No. 3.8e-73;
 Matches 836; Conservative 0; Mismatches 584; Indels 91; Gaps 6;

```

OY 547 gttatgaacatcatttttagttgattgctgcgaagtaaaaaagttggcattgtgtgtg 606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2082 gggactatcatgtatgtatgacacagcactgtcaaaaaaacggtatatactgttacc 2141

```

OY	607	tcocaaagctcgtgtccttctgtgtgctgttaaagctgaaacttgtttaaaactgtataaat	666
Db	2142	actctgtcaaaaggtgttggttgggtgtcaacaatgctcaactttaaacaacaactggt	2201
OY	667	gaatgaagaagatggaagaaagaatgttccgtgtcaactgagtttaactggttcttaa	726
Db	2202	atgattcaagaagacatgatattctataaagaatatgtcctaactgactatgtgtgtaca	2261
OY	727	ctctaatcttttccaacttcttggaacttggaataaattttaaccgaattcagaatttgt	786
Db	2262	ggggcggatcttctttaaagcagcgggaatcsgatgttccctatgactgaaggtctaacg	2321
OY	787	gatactccctcgtttggaagcagtttaattggttcagtcgaattcattgatgtatgtataaa	846
Db	2322	gattcaaccaagcttggaaagaagtcctgtaagatttggaaacggtgaactgtataccaa	2381
OY	847	aaaggggaagtagatgtgtgttattagtattaaacaaatttataacacagatgtgcgcaa	906
Db	2382	aacgagatgttngatgaattgattgtattgttataacccaagtttaacttcaactaagagt	2441
OY	907	aagcagatattggaataataatttcattaccagaatttagaataatgattgaattgaacgcaa	966
Db	2442	cagttctgtgtggaaaaaatgtgtccaattactgtatta-----gattccacgaagggc	2495
OY	967	agaaaaaaggttggagattatataatttaacgaactgtatgcgaaagtattatagaataatta	1026
Db	2496	acctcttaacgaacaagaatatctttagtggccttcaacctgaagcaatctgtgatacatta	2555
OY	1027	tgtgttcgttatttagaatctcgaagttatacagcagcagatgtgaaacctgtctttg	1086
Db	2556	ttaccacaatatgcagaagaagcttaatttattggtgcgattatgtatgcgaaacgaagaa	2615
OY	1087	caagcgcgcgcgaatggtgcgcacatgaagaagcaacgaactgaacgcagttactaattaat	1146
Db	2616	catgtcgtcgggatgtgcgcgcgaagaaacgaacgcacgcgaaacgaatgtataattatgcg	2675
OY	1147	gagttcacgttgatgcataacaagaagctgtcacaagcagattatacaaatgaataaatgaa	1206
Db	2676	gatttaacaatattctataaacgtgtcgtctcaagggcgattaccacaagaattacagaa	2735
OY	1207	atgtgtccggtgacgacgaatlttaacaat-----	1238
Db	2736	atgtgtcgtgtgcagctgcactagataatactagtcttaacgtttgaaagttataaaaa	2795
OY	1239	-----agaggtacgtgaattgtgcaactgtgaataat	1267
Db	2796	tgttcaaaagcgtgtgtgacttataatgaagactgvgaggaagaagaatatgatgttcagaaagat	2855
OY	1268	tgtcaacaatcatcgtgcgtgttataatgaactgtgattccccaagatgcaag--"taccaaa	1324
Db	2856	tgttcaagatcatcgtgtccggtttgttgacgtgagattttcatatgatcaatccttaaccga	2915
OY	1325	agtatatgatgccttaaatatgttgaacaagtt-----taagtacttgaa	1367
Db	2916	tatttaacaacgctttagctgttataaaaaatgvcgaaacaaacaaaagaatgattacttgaa	2975
OY	1368	gttcaacaacaatataggttgggtgtgttgatgttcgcgtgtatcgcgaatggatcatcgtatgaa	1427
Db	2976	gtcgcctttaaactaagatgcatgagatgagatattcgtctctatgctatggaatgtgcacagatggt	3035
OY	1428	ttaaaaagcgggttaaagcgttaacaacaatacgaataacccaattctgttccagttggaagc	1487
Db	3036	ttacaacgttggaatggaagtatacgtatacaggaataacatcatcttcagttccctgtgtgtaa	3095
OY	1488	aaaaactgggtcgtatcatcagaacgtatttgggtgaacccaatcgatgvcgaagtggaatc	1547
Db	3096	gatacatatgctgtcgtgtttaaagcgtttttaggaacacaatgtacttaagaagcgcaatc	3155
OY	1548	ggttcgaaa---aggaattgtctcattcacacgcgtgcgcacaaagtattgaaagaacaatc	1604
Db	3156	cctgcgaatctgtgaacgtatgagatcatcataaaaaagcgccagcagatttgatgaattagtt	3215
OY	1605	aacagttactgaacttttagaagaacgggaattaaagtatacgaactagttgttcggttcgcg	1664

[illegible]

RESULT 8

ID	AAF28551	standard; DNA; 94750 BP.
AAF28551		

AA AF28551; AC

DT 04-APR-2001 (first entry)

Genomic fragment #38.

Genomic library; bacteria; human upper airway; otitis media; sinusitis;

XX
XX
XX

[illegible][illegible]

XX 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040

XX	77	67	60	50	40	30	20	10	0
----	----	----	----	----	----	----	----	----	---

XX	Y	Z	W	V	U	T	S	R	Q	P	O	N	M	L	K	J	I	H	G	F	E	D	C	B	A
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26

XX

XX

XX
XX

compositions, and for identifying virulence factors, regulatory

PT acids -

PS Claim 1; Page 415-436; 545pp; English.

CC The present invention relates to a *Moraxella catarrhalis* genomic library

AAF28514-AAF28554). The library has a number of uses described in the

compositions. *M. catarrhalis* (Branhamella catarrhalis) is a large

flora of human upper airways. *M. catarrhalis* is known to cause acute, chronic sinusitis and bronchopulmonary

CC infection and life-threatening, systemic diseases including endocarditis

FT	CDS		complement (39873..40514)
FT		/tag=	k
FT		/label=	MG034
FT		/note=	"Previously identified as MORF-20101, the encoded protein shows 48.13 percentage identity to thymidylate kinase (tdk) from B. subtilis"
FT	CDS		40543..41787
FT		/tag=	l
FT		/label=	MG035
FT		/note=	"Previously identified as MORF-20102, the encoded protein shows 30.71 percentage identity to histidyl-tRNA synthetase (hss) from Mycobacterium leprae"
FT	CDS		complement (44751..46277)
FT		/tag=	m
FT		/label=	MG038
FT		/note=	"Previously identified as MORF-20105, the encoded protein shows 46.83 percentage identity to glycerol kinase (glpk) from E. coli"
FT	CDS		complement (46268..47422)
FT		/tag=	n
FT		/label=	MG039
FT		/note=	"Previously identified as MORF-19831 and MORF-20106, the encoded protein shows 43.20 percentage identity to glycerol-3-phosphate dehydrogenase (GUT2) from S. cerevisiae"
FT	CDS		49377..49643
FT		/tag=	o
FT		/label=	MG041
FT		/note=	"The encoded protein shows 48.86 percentage identity to phosphotransferase (ptsh) from Mycoplasma capricolum"
FT	CDS		50060..51520
FT		/tag=	p
FT		/label=	MG042
FT		/note=	"Previously identified as MORF-19832 and MORF-20108, the encoded protein shows 41.92 percentage identity to spermidine/putrescine transport ATP-binding protein (pta) from E. coli"
FT	CDS		51525..52382
FT		/tag=	q
FT		/label=	MG043
FT		/note=	"Previously identified as MORF-20110, the encoded protein shows 26.51 percentage identity to spermidine/putrescine transport system permease protein (ptaB) from E. coli"
FT	CDS		52366..53320
FT		/tag=	r
FT		/label=	MG044
FT		/note=	"Previously identified as MORF-20111, the encoded protein shows 29.45 percentage identity to spermidine/putrescine transport system permease protein C (ptaC) from E. coli"
FT	CDS		54658..55605
FT		/tag=	s
FT		/label=	MG046
FT		/note=	"Previously identified as MORF-20112, the encoded protein shows 36.60 percentage identity to sialoglycoprotease (gcp) from Pasteurella haemolytica"
FT	CDS		complement (56970..58310)
FT		/tag=	t
FT		/label=	MG048
FT		/note=	"Previously identified as MORF-19834, MORF-20114 and MORF-20115, the encoded protein shows 43.02 percentage identity to signal recognition particle protein (ffh) from B. subtilis"
FT	CDS		58117..59079

FT	/*tag=	u	
FT	/label=	MG049	
FT	/note=	"Previously identified as MORF-20114 and MORF-20115, the encoded protein shows 44.78 percentage identity to purine-nucleoside phosphorylase (deob) from <i>E. coli</i> "	
FT		59083..59754	
FT	/*tag=	v	
FT	/label=	MG050	
FT	/note=	"Previously identified as MORF-20117, the encoded protein shows 83.03 percentage identity to deoxyribose-phosphate aldolase (deoc) from <i>Mycoplasma pneumoniae</i> "	
FT		complement (64898..65731)	
FT	/*tag=	w	
FT	/label=	MG056	
FT	/note=	"Previously identified as MORF-20122, the encoded protein shows 30.25 percent identity to the protein disclosed in GB:D6185.99 from <i>B. subtilis</i> "	
FT		complement (65713..66249)	
FT	/*tag=	x	
FT	/label=	MG057	
FT	/note=	"Previously identified as MORF-20123, the encoded protein shows 38.90 percentage identity to the protein disclosed in GB:D26185_104 from <i>B. subtilis</i> "	
FT		81047..82597	
FT	/*tag=	y	
FT	/label=	MG067	
FT	/note=	"Previously identified as MORF-19845, the encoded protein shows 28.84 percentage identity to glutamic acid specific protease (stape) from <i>Staphylococcus aureus</i> "	
FT		91065..91919	
FT	/*tag=	z	
FT	/label=	MG070	
FT	/note=	"Previously identified as MORF-20136, the encoded protein shows 34.8 percentage identity to ribosomal protein S2 (rps2) from <i>Spirulina plantensis</i> "	
FT			

Query Match	15.1%	Score 297	DB 18	Length 580073
Best Local Similarity	49.6%	Pred. NLo. 4.5e-63		
Matches 979	Conservative	0	Mismatches 935	Indels 61
				Gaps 6
QY 10	ttggctcagtcgcgtcgtcgatcttctaagatcaacgtaagaacatcatcacaagtca	69		
Db 505474	TTTTACTCAATTTGGTAGTGATCTTGATGAATAATCAAAAAGGTTTTGAGCATTGTAA	505415		
QY 70	aaaagtaactgaattttgaacaaacacatacttcctcgtaactcgtgacacaaagc	129		
Db 505414	AGAGGTAAATGGAAATATTTAAACAAACCAATGGTAAACCTTACTCTCAATCCATGAAAC	505355		
QY 130	atgaagttaattgcagtagagtttgtaactctagaagaagtgaacttaagttatgg	189		
Db 505354	ATTATTTTATTGTCATTAACAACAACCTTCATTAAAGTTATTCAGTTGATGAATAATGC	505295		
QY 190	ttaacttgaatcagaacttttagagatgctaacacataactatgctgattttatgctga	249		
Db 505294	TAAATTTTAAACAAGATATACAGAAAGATTTAATGTTCCCAATCCTGTGTTAAAGATT	505235		
QY 250	gttaacccaatctgcgaacttaacaatgatga-----ttaaga	288		
Db 505234	ATCTAACAAAAAAGAAATTTACTAGAGATTTTAGAAAGTAAACATAAAACCGCTTTTAAAT	505175		
QY 289	gtcataaaagcatttttgatagctcaagaacaaacagtgcgtgtaagttaacactt	348		
Db 505174	GCCTTGAACGTTTTATCATGTCATTAACAGATTATGATATTACCAAAATTTGGTACTAT	505115		
QY 349	aaatgagagacaaatagcgaagtgcctaaagagataagaaacaaatcgcgagtgtaaa	408		
Db 505114	TGAGGAATTAATTAATGAGGCTTTTATCAAGAATAATTAAGCGCAATGAAATACAGTAAA	505055		

QY	409	agcacaacaaatctaataagcagatggaatggttgctgcccgcgaatatcgtaaacg	468
Db	505054	TCACACATTAGATACCTAATGCAATGAAATAGTGTCACGCCCTAAGTTATTAAAGTTC	504995
QY	469	caagaaacgcatgcttcctcaccgcccattccgaacacatacgaacgtaagtaagccac	528
Db	504994	AAAAAACAGTTTCAGAAATTAAGTTGTTGTTTTAAAGAAATTTAATAAGCTGTGGCCAA	504935
QY	529	gtttccaaagcaacgattggttacaagcatcatttttagtgcgcgaagtaaaaaa	588
Db	504934	GT-----AGTTGTTTTCTTTAAAGAACCAAAAAAGAACAGTAACCAAAAA	504887
QY	589	gtggcagatattgtgtccacagatcgcgtgctttgtgtgcttaaacgcygaactg	648
Db	504886	ACTTATAGATTAATGATGAGATGCTCTCTTTAGACATTGTGACAGCATTAATGCAATG	504827
QY	649	tttaaacitgataataatgaatgaaagaatgaaagaagaatgattccgcgtcaatg	708
Db	504826	AATAGCTTATTAATAAGCTAATTTTAAAGCGATGATGAATAATCTTTTTTATGCTAAGAAA	504767
QY	709	agtttaacgcttcatactatcatcaacttttcacatccttgggaattaaatttaacc	768
Db	504766	AACCAATCATTTCGAAATTAATAATAGTCATATTAATCCTGCTGTGGATTATTGATATC	504707
QY	769	caagattccgaattggtgtaactccctcctgttcgaagtaattggttcgaacttct	828
Db	504706	CAAGATCGTGATATT-----AATTTGATTAATGTCMAAGATATTGTGATCAG	504659
QY	829	atgattgatgctctataaaaaaggggaagatgagatggtgtaattagttataacaattt	888
Db	504658	ATTATGATGATGATTAAGATTAAGATTTAAAGCTTGATGCAATTTGTAATGCTTAACCTAAATTT	504599
QY	889	attacacgcatgctgcacaaagccagtactggaaaaaattatccatcacagatttagat	948
Db	504598	AAAAACCATTAATATCCAAACATGTCAGCTCTTTCAGATTCTTCTTCATGATTGGAACCT	504539
QY	949	aatgataaattagcggaagaaacaaagtttggattataatttcgaacctgtaagcga	1008
Db	504538	TTTTAAACCTTTAAATCCGGTTGTGATGATGATCAACCACTGATTTTACAGCCAGATCAAGCC	504479
QY	1009	gtattcttagataattatctatgctgtctatctagaatcctcagttatacaagaagaagt	1068
Db	504478	ACGATTAATTAATTAATTACTCCACAGTTTTCGATGATGCTCTTGATGATGCTGCTTTGTT	504419
QY	1069	gaaaaccttgcttcctgcgaacgcgcgtcgaaigtgctgcgaatgaagcagcaacaataac	1128
Db	504418	GAACCTAAGTTATGTCGATCAGCTTTCAGCAAAATGCAATGAGATGAGATGCTACAAAGAAAT	504359
QY	1129	gcaggttaacttaattatgtgttacagtttagtctatacaaaagctcgtcaagcaattat	1188
Db	504358	GCTAAAGATTACTGTGATTAATACACTTTTCAATTTAAACAGCTTAAGACAAAACCTGATTT	504299
QY	1189	acaaatgaatataatgaatctgtgtccggtgcgcgaacgaatatacaaaatagaggtcgg	1248
Db	504298	ACGAAACAGATTATTAATGAGATTATAGGAGGATGATGAATTAATTAATTAATAAAAAAGAAACCT	504239
QY	1249	taatgcaacttgaaaaattgtacaacaatcgcgtgcggtatttgacgcytgaattccac	1308
Db	504238	AA-----CATATGTTAAAGTTCCACCAAGTCATTTGGTCTCTGATGATGTTATCTTTTCAG	504183
QY	1309	aa-----ggtgcagtaacaaagtatagatgcttaaatgtg-----aaacggtt	1368
Db	504182	AAAGTAAACATTAATACCTAGAGATTATGATTTGTTGATGATGACAACTRAAAAAAAGTGAGC	504123
QY	1357	taagtaactgaagttcaacaacaatgaagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	1416
Db	504122	TTTTTTTAAAGCAACCACTAATTAATAGGTGATGACATTTGTTGCTGATTCGATTTGAGTTC	504063
QY	1417	catctatgattataaagcgcggtttaagcgtacaacaatacgaataacccaatttcgttc	1476
Db	504062	CTAACAAGAGGATTAATGACCGTAATGTTTAAAGTTTCTTAATTAACCATTCATTAAGAGGTGAC	504003

Oy	1477	caaggaggaacgaagaaacgttggctgcgtcctctgaaagcgtatgggtgaaacaaatcgatgagc	1536
Db	504002	CTGTGGCGAAAAAAGTATTGGGGAAGAGTGTTCATGTTTAAAGGTGAACCCATTGATGGA	503943
Oy	1537	aaggTgaatcvgTgTgcgaagagagattgTctatctcacgTgTgcacccaagtTatgaag	1596
Db	503942	AAGAAACCATTTCCAAAAAACCACAAAGCTATTCATCCATCGTAAACCAACCGCTTTTGATG	503883
Oy	1557	aacaatctaacagtaactgaaacttttagaaacgggaatlaaagttatccagactatgcttgc	1656
Db	503882	AACACACAAATACCTGTGAATATTTTCAACACAGAAATAAAGTAATGACTTTTAATC	503823
Oy	1657	cgttTgcgaaggggTgaagtagttatctgcgtTgTgcggTgTgcgTgaaaacgTca	1716
Db	503822	CTTACGTTTAAAGGGGGGCTTAAATGCTTATTTTGGAGAGCGCTGGTGTGTGGAAACAGCTTT	503763
Oy	1717	atatgatgaaatlaataccgtcaatccgaaatTgagcaactcaggttaactctgtctttgcg	1776
Db	503762	TGGTGCACGAAGATTTTATTTCAATACACTTCCCAAGAACACTTCTGGTTTAAAGTATTTGCTG	503703
Oy	1777	gggtgagTgagcgtgacgTgcgaagtagtaagactctatctctgagatgaaagactctaacg	1836
Db	503702	GAGTTTGGTGAAGAACAAAGACAAAGAGTAAATGATCTTTTACTTGAATGATTTCAAGGTGGGG	503643
Oy	1837	tattagaataagTgctctctgtttatTgtccaatTgaacagcgcacacagTtaacgctttac	1896
Db	503642	TGATGATTAAAAACAGTTTATGTTTGGTTCGCAAAATGAAATGAACCAACGAGCTAAGAAATGA	503583
Oy	1857	gtgtgTgcattaacagcgttaactatTgTgcgTgaaaaaatctcgtTgatgaagTcgtTga	1951
Db	503582	GAGTGTGCTTTAAACGTCTTTAAACAATGGCAAAATATTTTCTGTGATCATGATAAATCA	503528

RESULT 10
AAC79635
ID AAC79635 standard; DNA; 531 BP.
XX
XX
AC AAC79635;
XX
DT 08-FEB-2001 (first entry)
XX
DE Virulence gene #44.
XX
KW Virulence gene; antibacterial; vaccine; bacterial infection;
KM septicemia; bronchopneumonia; rhinitis; wound infection; ss.
XX
OS Actinobacillus pleuropneumoniae.
XX
PN W0200061724-A2.
XX
PD 19-OCT-2000.
XX
XX
PF 06-APR-2000; 2000MO-US09218.
XX
PR 09-APR-1999; 99US-0128689.
PR 10-SEP-1999; 99US-0153453.
XX
XX
PA (PHAA) PHARMACIA & UPJOHN INC.
XX
PI Lowery DE, Fuller TE, Kennedy MJ;
XX
XX WPI; 2000-647422/62.
XX
DR
XX
PT Attenuated Pasteurellaceae bacteria comprising mutations in virulence
PT genes, useful as a live attenuated vaccine against bacterial infections
PT
XX
PS Example 5; Pages 238-239; 322pp; English.
XX
XX
CC The family Pasteurellaceae encompasses several pathogens that infect a
CC wide variety of animals. The present invention relates to virulence genes
CC from Pasteurellaceae. The present sequence is one such virulence gene.
CC The present sequence may be mutated in order to produce an inactive gene.
CC

QY 491 gccctattcagaacaacgtaacgtat-----tagccagctttccaagcaaga 544
 DB 3608 aagttacgtcagaagaagtcgttaaacctttgacagataccttccatgtaacgtgagctg 3667
 QY 545 ttggttacgaagcatcatttttagtgatcgcaagtaaaaaaagcgggacgtatggttg 604
 DB 3668 ggtgttcaactaactccgagtggtgattagccgtctctgtaagaagacagctatatacgtta 3727
 QY 605 tgtccacagatcgtgtgcttctgtgtgcttaaacgttgaactgttttaaacgttataa 664
 DB 3728 tcaactcagaccggtgttgggtggtgattatcctcctattttgaaagcgttatg 3787
 QY 665 atgaaatgaagaagatggaagaaagatggttcogttcaattgagtttaactggtctta 724
 DB 3788 agttgaaagaagatlaaccccccagacgtgaaggttttgaatgatactgtatcggtgga 3847
 QY 725 aatctacaacttttccaactcttggaattaaattttaaccagaattcagatg 784
 DB 3848 tgggagctgatacttcttaaggtcgcggtatccaactttatgatactgagctgtg 3907
 QY 785 gtgatacctcctgtgtgacgttaattgtgtcagtaactctatgattgagcttata 844
 DB 3908 ccagaccactagcttggatccaagttcgttaagattatttcaaaaactgtgaaatgtaac 3967
 QY 845 aaaaaggggaagatggtgtgtattgtattatatacaaatattataacagatgtgc 904
 DB 3968 aaaaatgaactcttgaagcttattgtttgttgcataccaccatgtaacgttaacca 4027
 QY 905 aaaaagcagatltggaanaatttaattccattaccagaattatgataatgaaattgagc 964
 DB 4028 gtcaaatgtgtggaacaaatgcttcgattgtgtgacttgatgtaacaaatga---agcgg 4084
 QY 965 aaagaaacaagtttggatattatattacgaacctgtatgcgaagatattatagataat 1024
 DB 4085 atgaagaagtaacagcttgactttgaaattgaaaccagcagaagaatcttgagcagt 4144
 QY 1025 tatgtgtgtattattagatctcagtttatataacagcagttgaaacctgtcttg 1084
 DB 4145 tctgtcctcagtttgcagaagaatgattatcaggttccattacagatgccaagacagct 4204
 QY 1085 agcaaacgcgtcgaatggtcgcacatgaaagcaagcaagataacagcaggttaattat 1144
 DB 4205 agaatgtcgtggcagatgacacagccaatgcaacagcagataatgcttaagaagatca 4264
 QY 1145 atgattacagttagttatataacaaagctcgttcaagcaagattatacaaatgaattaat 1204
 DB 4265 atgatttgcaattcagttataacgtgcagacagcagcagattacacaagaattacag 4324
 QY 1205 aattgttgcgtgtgacagcaaatlaacaatagagatc----- 1246
 DB 4325 aaatcgtagcaggtgctagttccttagaataagctctagctcgtatgaaatga 4384
 QY 1247 -----ggtaatg 1253
 DB 4385 cttagagactagtgtgacgacgaaccgacgtatcttataatagaaagagagatg 4444
 QY 1254 gcaactggaanaatttgaacaaatcaggtgtggttattatgacgttaattccacaagaat 1313
 DB 4445 agttcaggttaaatgtcaggttaccgttccggtgtgtagagctttgtttgacagag 4504
 QY 1314 gcaatgcaaaaagtataatgcttaaatgtg-----aaaca 1352
 DB 4505 gaaaacactcttcgtagattacaatgacattgctgtctacaaaaatgacgaagaanaa 4564
 QY 1353 ggtttagtactgaagttcaacaacattaggtgtgtgtggtatgctgtatcgaatg 1412
 DB 4565 aaaaatgcttctgaagtagccttggttagagatgagatgattgttctgactatccgcaag 4624
 QY 1413 gatactctgattgataaagcgggtttaaagcgtlaacaatacgaatacccaattct 1472
 DB 4625 gatacaacagatggtgtgactgtgtgaatggaagtatttgacacaggtctgtccaatctc 4684

QY 1473 gttcaggtgggaacgaanaacattggtgtgtatcatgaacgtattgttgaaacacgtat 1532
 DB 4685 gtaccagtaggttaagaacatttggagcgtgtcttcaacgcttttggagatatactgac 4744
 QY 1533 gaaaggtgaanaatcgtgtcaga---agaaattgttctatccacgttcgcaccaaat 1589
 DB 4745 ttggaagctccttttcaagaagcagagcagatgcacaaattataaaaaagttccaact 4804
 QY 1590 tatgaagaacatacgaactgaactgaacttttagaaacgggaattaaattatcgactta 1649
 DB 4805 ttgtagatgtgtctaaccttcttgaatcccttgaaacagggaatcaagttattgacctt 4864
 QY 1650 gttgtccgttttgcgaaggggttaagtagttatttgcgtgtgtcgtgtgtgaa 1709
 DB 4865 ctgccccttaaccttaaggtgtgaagttggaacttttcgtgtgtgcggaggtgtgaa 4924
 QY 1710 accgtcaataatgataatgaatcgcataacatcgcaattggaactcagatcgtctgc 1769
 DB 4925 actgtcttaatccaagaattgtatccaacatgcccagaagcaggtgttatttca 4984
 QY 1770 ttgcgggggtaggtgagcgtacggtgaaggtaacgacttataatgagatgaagaac 1829
 DB 4985 ttgtcgtgtgtgtggaacgtactcgttgagggaatgaccttactcgtggaatgaaga 5044
 QY 1830 tctaaagtattagataaagtgtctctgtttatgtgtcaaatgaaagcagcaggttaac 1889
 DB 5045 tcaagcgttatacgaanaaacagcagatgcttctgttcagatgaatgagccacagaaga 5104
 QY 1890 cgttaccgtgtgcaataacagcgttaactatgctgcggaanaatccgttgatgaagtcgt 1949
 DB 5105 cgtatgcgtgtgccccttactgtttgcaaatcgtcgaatactccgtgtatgtggaagc 5164
 QY 1950 gatgtctattctcgtt 1967
 DB 5165 caagacgtgtctctctt 5182

RESULT: 12
 AAV29571
 ID AAV29571 standard; DNA; 4815 BP.
 AC AAV29571;
 XX
 DT 04-AUG-1998 (first entry)
 XX
 DE L. lactis soluble part (F1) exhibiting ATPase activity encoding DNA
 XX
 KW Biomass production; uncoupled ATPase; F0F1 ATPase; membrane bound;
 KM F1; Lactococcus lactis; ds.
 XX
 OS Lactococcus lactis.
 XX
 FH Key
 FT 26..553 Location/Qualifiers
 CDS
 FT
 FT /tag= a "ATPase subunit".
 FT /product= "ATPase subunit".
 FT /gene= "atpH"
 FT /note= "delta subunit of the F1 portion of F0F1 ATPase"
 CDS
 FT 742..2244
 FT /tag= b
 FT /product= "ATPase subunit"
 FT /gene= "atpA"
 FT /note= "alpha subunit of the F1 portion of F0F1 ATPase"
 CDS
 FT 2260..3129
 FT /tag= c
 FT /product= "ATPase subunit"
 FT /gene= "atpC"
 FT /note= "gamma subunit of the F1 portion of F0F1 ATPase"
 CDS
 FT 3301..4710
 FT /tag= d
 FT /product= "ATPase subunit"
 FT /gene= "atpD"
 FT /note= "beta subunit of the F1 portion of F0F1 ATPase"

XX WO9810089-A1.
 PN 12-MAR-1998.
 XX 08-SEP-1997; 97MO-DK00373.
 XX 06-SEP-1996; 96DK-0000963.
 XX (JENSEN/ JENSEN P. R.
 XX Snoep JL, Westerhoff HV;
 PI WPI; 1998-193637/17.
 DR P-PSDB; AAM56790, AAM56791, AAM56792, AAM56793.
 XX PT Method improving production of biomass or a desired product - by
 PT expressing an uncoupled ATPase activity in the cell
 PS Claim 16; Pages 35-41; 78pp; English.

XX This DNA encodes the soluble part (F1) of membrane bound (F0F1 type) H+
 CC - ATPase or a portion of F1 exhibiting ATPase activity. The DNA is
 CC derived from *Lactococcus lactis* subsp. *cremoris* strain MG1363. This is
 CC used in a novel method for improving the production of biomass or a
 CC desired product from a cell. The method comprises expressing an uncoupled
 CC ATPase activity in the cell to induce conversion of ATP to ADP without
 CC primary effects on other cellular metabolites or functions and incubating
 CC the cell with a suitable substrate to produce the biomass or product. The
 CC expression is directed using a vector including DNA encoding the soluble
 CC part (F1) of the membrane bound (F0F1 type) H+-ATPase or a portion of F1
 CC exhibiting ATPase activity, the DNA being derived from *Lactococcus lactis*
 CC subsp. *cremoris*, *Lactococcus lactis* subsp. *lactis*, *Streptococcus*
 CC *thermophilus*, *Phaffia rhodozyma* or *Trichoderma reesei*, where the DNA is
 CC under the control of a promoter. An ideal ATPase is the membrane bound
 CC H+ ATPase. This enzyme complex consists of two parts, the membrane
 CC integral part and the (F0) and the cytoplasmic part (F1). Together the
 CC two parts couple the hydrolysis of ATP and ADP to the translocation of
 CC protons across the cytoplasmic membrane, or vice versa. The proton
 CC gradient is used to drive ATP synthesis from ADP and Pi. The method can
 CC be used for optimising the formation of biomass or a desired product,
 CC e.g. the product may be lactic acid which results in the acidification of
 CC dairy products. The method is more efficient than currently used methods
 CC of biomass production.

XX Sequence 4815 BP; 1491 A; 897 C; 981 G; 1446 T; 0 other;

Query Match 13.28; Score 261.2; DB: 19; Length 4815;
 Best Local Similarity 50.48; Pred. No. 6.2e-55;
 Matches 1061; Conservative 0; Mismatches 838; Indels 207; Gaps 9;

QY 54 aattatccatcgtgcaaaaagtaactgaattatgaacaacaacacatctccgtat 113
 DB 1967 aattgaatcgtggtcgtcgtcgtcgttgaagtcctgaacaacacatcgaacaacattggtg 2026
 QY 114 ctgtgacacaacaagcattatgtatttgcagtagagtttggttacttactaagaagttg 173
 DB 2027 ctgtgaaacaagaatttgcattctctatgcatcattgcatcattgtaaatgttc 2086
 QY 174 actagatcgtatgttgcatttgaatcagaacctttagatgatgtcactaactatg 233
 DB 2087 cagtgtgatgttcttcttatttgaactaaatgttcgattcttcgattgcaattatg 2146
 QY 234 ctgatttattcgtgtgagtttaaccatctgccaattacatgatgaataaataagtgat 293
 DB 2147 cagatcctcttgaaagtaactactgacacaaagatttccagaagaagcaaaacttgagc 2206
 QY 294 taaagggcatttggatagcttcaagaacagatgctgtgtaagttaacatttaag 353
 DB 2207 ---aagcaattaaagcattcaaaaatacacagcaattatttaagaaggagcttaactaag 2263
 QY 354 gagagacaanaatgtagtgcttaagaagataagaacaaatcgcgagtgttaaaaaagtac 413

DB 2264 gaggcttcaact-----taacgaataaanaactaagatgcttcaacaagaagaac 2312
 QY 414 acaaaaattactaaagcagatggaatgttctgtcctcgtgaanaatgcyfaaaacgcaaga 473
 DB 2313 aagtcacaatcagcagtgatcgaatgaatgttctgtcgtgaacttcaaaagcagaatc 2372
 QY 474 agcagatcctcttcaacgccttattcagaacaacatcagtaacgttattagccac---gt 530
 DB 2373 tcacgctaaagccttccagacttattcgtgaanaagtaagtaagattacgactgacttagt 2432
 QY 531 ttcaagaacagcagatttgttaccagatccatttttagttagtgcgcaagtaaaaaagt 590
 DB 2433 ttcaagcagataatgagccggcccaaaaatccgatgatgatgataacgtagcaagaaac 2492
 QY 591 gggcagatgattgtgttccacagatcgtggtccttctgtgtggttaaacgttgaacttgt 650
 DB 2493 tggctatcctgtatcacaacagatcgtggtgttctgtgtgcaatttaaatcaaatltt 2552
 QY 651 taaactgtatttaaatgaaatgaaatgaaagaagaagaatggttcgttcaattgag 710
 DB 2553 gaagctgttataagtaataatcagtaaacgcacacaacaatgagagtgagatatacaact 2612
 QY 711 ttaatcgttcaataatcatacactttccaatcttcttggaattaaatlttaacca 770
 DB 2613 tgccttgtgtgtaacgagcgagacttttcaaaagcccgtaacgtaaaagttccttattgt 2672
 QY 771 agatcagattgttgtatccctcctgttgagcagtaattgttctagtaactat 830
 DB 2673 tctcggagacttcagataacacgaccccttgaagagttcggcgcaattgttcaagaagc 2732
 QY 831 gatgatgcttaaaaaaagggaagtagatgttgttattttagttatatacaaatltat 890
 DB 2733 cgtgaagaataatcaacagaagaatcagatgtaactttagttgttatacaacaacatgt 2792
 QY 891 taacagcatgtcgaaaaagcagatgttgaanaaatlaattccattaccagaattagtaa 950
 DB 2793 gaactcattgttaagtgaagcagcagatgagaanaaat-----gttacctatttcttga 2846
 QY 951 tgtatgaattgagcgaagaagaacagtttggattatatttaacgaactgatgsgaagt 1010
 DB 2847 tgaanaagtgtagaanaaagcatctctgtacatttgaatttagaacagatcgtgaac 2906
 QY 1011 attttagtaattatttgtgttcttatttgaatcctcaggtttatcaacgaagttga 1070
 DB 2907 aatcctaactcagttgttgcgcgacatatgtcgaagaatgattatgttgcctcaattgtga 2966
 QY 1071 aaacttgccttgcagcaagcgcgtcgaatgttgcacatgaaagcagaacaacatagcgc 1130
 DB 2967 tgcanaaacagcagagaacatgtcgtcagatgacgcgaatgctgactgcacaagaataatgc 3026
 QY 1131 aggttaacttaattatgagttacagttcagttataacaagaactgttgaagaagtattac 1190
 DB 3027 acattcgtcattatgatttaacatcacaatlaaacgttctgcgaagcttcaattac 3086
 QY 1191 aatgaatttaataaatttgcgtgcgtgcagcagc----- 1226
 DB 3087 gcaagaaattcaaggaattgttgcggtgtcctcagcgtataactgttcaaacattat 3146
 QY 1227 ----- 1226
 DB 3147 tctcaatgttaacgttatacacttgaggaataaatgttctcgtcagtaagaagcttgaat 3206
 QY 1227 ----- 1226
 DB 3207 tttaatacgtttgtcagtaaatlttlaactgattagctttaanaatgaaatagaaattctgt 3266
 QY 1227 -----aatttaanaaatagagatcogttaaattgcaactggaanaattgttaacaat 1276
 DB 3267 tgttagacagaanaataaacaagaagggaanaaacatttgcgttctgttataactacaggt 3326
 QY 1277 catcgttgcgtttatttgcgttgaatttccacaagaatgc---agtaaccaaaagtatga 1333


```
Db 3327 tatcgctcccgctgctgagctggaatttggttctgtatgcacaactgcctgagattacaac 3386
QY 1334 tgccttaaty-----tgaacaagtttactgtgaattca 1372
FT ||||| | |
Db 3387 tgccttgattgtctacaagaatgtlcaatgtttaaacaanaattactcttgaagtgc 3446
FT ||||| | |
QY 1373 acaacattagggggggtgtgtcgtatcgcaattgggggtactctgattgattaa 1432
Db 3447 ttggaactgtgtagtggtcgtatcgtaacgtcctatggaactcactgattgattgac 3506
QY 1433 acgagtttaagcgtlacaacaatgcgaataacccaatttctgtccagtgggaacgaac 1492
Db 3507 tctgtgacttgagtccttgatcacagtgtaaggggtccagcgttccctgttggtgactac 3566
QY 1493 attgggtcatcactgaaacgttgggtgaacccaatcgatggaaggtgaaatcggtgc 1552
Db 3567 tcttgctgctgttttaattgctccttggtgacgttattgattggtggaagatttccctgc 3626
QY 1553 aga---aggaattgtgtatcattcacccgtgcgcacaagttaagaacaatacgaac 1609
Db 3627 tgatgagagacglaacccaatcccaacaagaagaagctcacaacttgcagaaattgcaactgc 3686
QY 1610 tactgaactttagaacacgggaattaaagttatcgacttaattgtccgtttgcgaag 1669
Db 3687 aatgagagctctgtacacagggataaagtgtcgatttactgtcccttacttacttaag 3746
QY 1670 gggtaagtagtttactcgtgtggtgcgttcgtgtaaaacccgtlcaatatgattgaat 1729
Db 3747 tgggaagtcggaactccttcgtgtgcgtgtgtgtaaaacccgtlcaatatgattgaat 3806
QY 1730 aatcgtaacatcgcaattgagagcactcagcttactgtctgtggtgggtgattgag 1789
Db 3807 gatccaataatgcacaagaacacgggtggtatcttctgtatccaaggtgtgtgcacgc 3866
QY 1790 tacgcgtgaagtgaaacgacttctatcatgagatgaagaagctcacaagttatgaataagt 1849
Db 3867 tactcgtgcgggaatgacttacttactgcgtgaagaatgaaagaatcagcgttatttgaanaac 3926
QY 1850 gtctctgttattgttcaaatgaatgaacgagccacaggttaacggttctggtgtgcattac 1909
Db 3927 agccatggtcttgtgtcaaatgaaatgaacacacgtgagacggtatgctgttgccttac 3986
QY 1910 aggccttaactgacgggaanaattcgttattg---aaggtcgtgattgttacttctgt 1966
Db 3987 tgggttaacaatgcggaatatttccgtgtatgttcaaggaacgaactgttcttcat 4046
QY 1967 tgataa 1972
Db 4047 cgataa 4052

RESULT 13
AAK29857
ID AAK29857 standard; DNA; 4001 BP.
AC
AC AAK29857;
XX
XX 22-JUN-1999 (first entry)
DT
XX
XX Thermotoga neapolitana type F ATPase gene.
DE
XX
XX Heat resistant; type F ATPase; enzyme; bioreactor; ds.
KM
XX Thermotoga neapolitana.
OS
XX
XX Key Location/Qualifiers
FH 130..1641
FT CDS
FT /tag= a
FT /product= "subunit #1"
FT /transl_except= (pos: 130..132, aa: Met)
FT CDS
FT /tag= b
FT /product= "subunit #2"
```

```
FT /note= "CDS appears to be missing 48 bases between
FT nucleotides 2040-2041"
FT CDS 2448..3806
FT /tag= c
FT /product= "subunit #3"
FT /note= "CDS appears to be missing 48 bases between
FT nucleotide 3386-3387"
PN JP11075855-A.
PD
XX 23-MAR-1999.
XX
XX 16-SEP-1997; 97JP-0250926.
PF
XX 16-SEP-1997; 97JP-0250926.
PR
XX
XX PA (KAIYO-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.
XX
XX MPI: 1999-257703/22.
DR P-PSDB; AAY02647, AAY02648, AAY02649.
XX
XX Heat resistant type F ATPase - useful as a heat resistant reagent
PT and an ATP-supplying enzyme in a bioreactor
XX
XX PS Disclosure; Page 9-13; 14pp; Japanese.
XX
XX CC This sequence represents a novel Thermotoga neapolitana heat resistant
XX type F ATPase gene which encodes 3 subunits of the protein complex.
XX CC The new protein is useful as a heat resistant reagent and an
XX ATP-supplying enzyme in a bioreactor.
XX
XX SQ Sequence 4001 BP; 1185 A; 872 C; 1095 G; 849 T; 0 other;

Query Match 12.7%; Score 250.8; DB 20; Length 4001;
Best Local Similarity 57.6%; Pred. No. 2,2e-52;
Matches 528; Conservative 0; Mismatches 352; Indels 36; Gaps 3;

QY 1093 gctcgaatgtgcgcattgaagaacgaacagataacgcaggttaacttaattgaattga 1152
Db 2281 gccagacagaaccccatgaaagaacgacgaacacacgccaagaaattgatacgaagagctg 2340
QY 1153 cagttgcttatatacaaatcgtcacaagaatgataatgataatgataatgattgtt 1212
Db 2341 acaactgcttatacaacaagaacgagtcaggttgcattacaccagaactctatgagatagt 2400
QY 1213 gccggtgcagcagaatttaacaatatagagatcggtaatg----- 1254
Db 2401 acgggtgcctgaagcactggaaggaatcgaaatatgagtgatgcgtatgcaaaaggtt 2460
QY 1255 caactggaaaaattgtacaataatcgtgtgcgtttattgtacgttgaattccacaagaatg 1314
Db 2461 caaaagatacatagtggaagtgatggtggcccgctgtgagatgaaatcccgagaag 2520
QY 1315 cagtaaccaagaatatagtatgcttaacttaactgaacacagttta-----g 1359
Db 2521 aactccaagacatatcaacgaccttgaggtcgtgatacctcaaaagggcggaagaatg 2580
QY 1360 tacttgaattcaacaacaattagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1419
Db 2581 ttcttgaagtagaacaacatcagtagagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2640
QY 1420 ctgattgattaaacggtttaaagcgtlacaacaatacgataaaccatttctgtccag 1479
Db 2641 ccgattggtcatalgaaggtgtctggaagtgtgtatcaggtgtgaacaaattacgcgtcccg 2700
QY 1480 tgggaacgaaaacatggtgtgtatcatgaacgtatgtgtgtgtgtgtgtgtgtgtgtgtgt 1539
Db 2701 tgggaagaagaatgatactcggaaagaatcgtgaacgttataaggagccgtgtgaagaacgc 2760
QY 1540 gtgaatacgtgtcagaagaatgtgtctatccacgtgtccacccaagttaagaagac 1599
Db 2761 gcgaagtaagtcgaagaagaaggtgtgcgcatccacagcgtctgtctgtactatagaac 2820
```

KW angiogenesis inhibitor; abnormal wound healing; inflammation;
 KW rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
 KW diabetic retinopathy; macula degeneration; haemangioma; detection;
 KW arterial-venous malformation; immune deficiency; ss.

OS Homo sapiens.

PN W0937660-A1.

PD 29-JUL-1999.

PF 22-JAN-1999; 99WO-0501313.

PR 23-JAN-1998; 98US-0072298.

PR 28-AUG-1998; 98US-0098539.

XX (IRUE/) IRUELA-ARISPE L.

PA (HAST/) HASTINGS G A.

PA (RUBE/) RUBEN S M.

PI Irue-la-Arispe L, Hastings GA, Ruben SM.

DR WPI; 1999-590684/50.

PT New isolated metalloprotease thrombospondin polypeptides, useful for
 treating hyperproliferative disorders, cancers or autoimmune disorders

PS Disclosure; Page 347-352; 457pp; English.

CC AA232000 and AA232001 encode, and AA49501 and AA49502 represent, human
 CC metalloprotease thrombospondin (METH) proteins METH1 and METH2
 CC respectively. METH1 and METH2 have been found to be potent inhibitors of
 CC angiogenesis both in vitro and in vivo. They can be used for treating
 CC cancer and other disorders related to angiogenesis including abnormal
 CC wound healing, inflammation, rheumatoid arthritis, psoriasis,
 CC endometrial bleeding disorders, diabetic retinopathy, some forms of
 CC macula degeneration, hemangiomas, and arterial-venous malformations.
 CC They may be useful in treating deficiencies or disorders of the immune
 CC system, by activating or inhibiting the proliferation, differentiation,
 CC or mobilization (chemotaxis) of immune cells. The etiology of these
 CC immune deficiencies or disorders may be genetic, somatic, such as
 CC cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
 CC toxins), or infectious. They can also be used to treat inflammatory or
 CC conditions, both chronic and acute conditions. The products can also be
 CC used for detection and diagnosis. AA232002 to AA232080, and AA49503 to
 CC AA49511 represent sequences given in the exemplification of the present
 CC invention.

XX Sequence 8008 BP; 1430 A; 2591 C; 2649 G; 1338 T; 0 other;

Query Match 10.6%; Score 210; DB 20; Length 8008;

Best Local Similarity 53.6%; Pred. No. 3.3e-42;

Matches 548; Conservative 0; Mismatches 440; Indels 34; Gaps 4;

QY 981 ggattatatattacgaactgacgataagataattatgacgattatt 1040
 DB 5250 ggtttacgactacgacgagcagatcgaatacctgacgagctgctccgctgagt 5309
 QY 1041 agaatcaggtttacgacgagcaggttgaataacctgtgtgacgacgagcgtgagt 1100
 DB 5310 ggcacagcgacgttgcgcgctgctggaatacaacgctcccttaacgaggcgagat 5369
 QY 1101 ggtcgcacgaaagcagcagacgataacgacgaaacttaacttaagttacgattagt 1160
 DB 5370 gtcgacacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 5429
 QY 1161 ctatacaagctcgtcaagcagatattacaatgaatgaatgtgtgcggtgc 1220
 DB 5430 gataaacgctcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 5489
 QY 1221 agcagcaatttaacaaa---tagagatcggttaatgca---actgaaaaaatgttaca 1273

DB 5490 cgaagcgtcttcaagcgaacccgagatagaagaaatggaacgaagcaagcaagtgaccaca 5549
 QY 1274 aatcatcggtgctgattatgacgttgattccaccacaaga-----tgaatgacaaaagt 1327
 DB 5550 ggtcattcgagcgtctgctgacgtgacgttgacgaagcgcctcccgacattctgacgc 5609
 QY 1328 atatgacgttaaatgttgaacaggttgaactgaactgaactgaactgaactgaactgaact 1387
 DB 5610 ccttgaacacccaacaaacgaacgacgacgctgttctcgaagtgagcagcagcagcagcag 5669
 QY 1388 tgggtgagtcgtctatcgaatggatcagatcagatgagatgaataaagcgcgttgaagcgt 1447
 DB 5670 gaacacgcgtcgcacacatcgatgacgacgacgaggggtctcgtgcgagcgcgagcgt 5729
 QY 1448 aacaaatacgaataacccaattctgttccatgttggaagaaacatttgggtcgtatcat 1507
 DB 5730 gtcgacacccgagcgcgcgtatccggttcggttggaacgaccccttggcgcagatcct 5789
 QY 1508 gaacgtattgggtgaacccaatcgatgagcaaggtgaatcggtgcagaagagattgctc 1567
 DB 5790 gaacgtacatcgagcagcgcgttggaacgacgagcgtgacgtgacgaagccgaagccgggc 5849
 QY 1568 tattaccgtgcgccaacgaatgaagaacacatcaacatcaacgtactgaacttgaagaac 1627
 DB 5850 gatccacacgcccgcgcgcgttctgcgcgcagtcgacgaggaagccagatccctcgtcac 5909
 QY 1628 ggaataaagttacgactgattgttcggttcggaagggggtaagattatt 1687
 DB 5910 cggcatcaagtgatgacgtctgcgcctccatccaaaggcgcgaagatcgtctct 5969
 QY 1688 cgtgtgtgaggttgcgtaaaacgctcaatgatgatgaataacccgaatccgacatccgaat 1747
 DB 5970 cggcggcgcgggttgggaacgctctgacatgaagaacgaacacacacacacgcgaa 6029
 QY 1748 tgaacacgaattactcgtcttgcgggggtgagtgagcgtacgctgaagtaacga 1807
 DB 6030 agtgcacgtggtctctcgtgttgcggggcgttgcgacgacgacccggtgagggcaacga 6089
 QY 1808 ctctatacagatgaagaacttaacgtattagat-----aa 1846
 DB 6090 cctttacacagatgatcgatcgagcgttatacaactcgaagcgtcgaagaatcgaa 6149
 QY 1847 agtgcctctgtttatgtgcaatgaagcagcagcagcagcagcagcagcagcagcagcagc 1906
 DB 6150 agtgcgctggtctcgcgcagatgaacaccccggggcccgtgcgcgcgtgcgcgt 6209
 QY 1907 aacagcctaactatggcgaagaatctcgttgatgaagtcgtgatgttcttctcgt 1966
 DB 6210 gacgcgcctgaaccccgcggaacagttccgcgacacgagtcggcagcagcagcagcagcag 6269
 QY 1967 tg 1968
 DB 6270 cg 6271

Search completed: October 28, 2001, 17:20:40
 Job time: 8409 sec

gene 4537..5007
/gene="atpf"
/note="PM1490"
4537..5007
/gene="atpf"
/codon_start=1
/transl_table=1
/product="atpf"
/protein_id="AAK03574.1"
/db_xref="GI:12721874"
/translation="MNLNATLIGQLIAFIAFVAFCKFWPPIIKAEERORSANAL
ASAARKEADKFTALVEQETFAKMOQIIDLAKRRELEVEKAEATKII
EOGAEEVEERKRVQELRVKVASLAINGAKRYGRVDEAANDIIDKIVAEI"
5021..5554
/gene="atpf"
/note="PM1491"
5021..5554
/gene="atpf"
/codon_start=1
/transl_table=1
/product="atpf"
/protein_id="AAK03575.1"
/db_xref="GI:12721875"
/translation="MSELTITARPYPAKAFDPAIBONAVEKMANLHFSSELKDETM
OTPKSSLASAKIADVTISIGCELDQYGOMLRLMAENKRELEVPATYOAFOHYVEE
HOKAEEVOYISAOPLNATQOEOKIASAMEKRLARKYKLCSDLSLIAVYIIRTDPMVI
DSSRGQLSRLANELQL"
5567..7108
/gene="atpa"
/note="PM1492"
5567..7108
/gene="atpa"
/codon_start=1
/transl_table=1
/product="atpa"
/protein_id="AAK03576.1"
/db_xref="GI:12721876"
/translation="WQNLSTELISELIKRIKIAQFDVYSARNTGTVSVSDGIRIHLG
SEVMQEMTALPGNRYAMALNLERDSVAVMGPADIAGEVOCETRIILEVGRG
LLGRVNTLIGPIDEKGEINDGSPGVIAAGVIDRKSVDQPVOTGKAVDSMPIG
RGORBLITIGDROGTALAIIDAIINQDSGKICIVAIAGKASTANVYRKLEESAL
ONTIYVVASASESALQYLAIPYSGCAMEGEPRDREGDALITIDLSKAVAVRISAL
LRPGRFAPYNGDVFLYLSRLERASRYNEVEPTGVEVKGSGTALPIPTOA
GDVAFVPTNYISITDQIFLESNLFNGCIRAPVPGISYRVGSSAOTKTKIAGG
IRIALAQRELAAFAQFASDLDDATRKQLSGQKYTELLKOKYSPLSVAQOALFA
VEFGLVEVDLDRIGSFESALLEYANHYADFMELTQSGYNDIEIKESLGIIDSFK
ANSAM"
7135..8004
/gene="atpg"
/note="PM1493"
7135..8004
/gene="atpg"
/codon_start=1
/transl_table=1
/product="atpg"
/protein_id="AAK03577.1"
/db_xref="GI:12721877"
/translation="MAGAKELRTKIASVSTOKITKANEMVAAKMRKTOERMSSRP
YSETIRNVISHVSKATIGKHPFLVDREVKKGMIVYSTDGKLCGLNVLKFLYANE
MKEMKEKVSQVLSIGSKSINFPQSILKITIDSGIDGPPSVQOLGVSNDMIDAY
KKGEVDVYLYINKEPINTMSOKPYLEKILPIELDNDLGERKQWMDIYERPAVLL
DNLVRYLESQYQAAVENLASBOARNAVAKAATDNGNLINELQVLVYKRAKASTI
NELVEIVAGAAAI"
8022..9395
/gene="atpd"
/note="PM1494"
8022..9395
/gene="atpd"
/codon_start=1
/transl_table=1
/product="atpd"
/protein_id="AAK03578.1"
/db_xref="GI:12721878"

gene 4537..5007
/translation="MATGKIYOIIGAVIDVEFPDAPVKVYDALNVTGLVLEYVOOL
GGGVRCJAMGSSQSLKRGSLSTNNTNISPVGKTKIGRLMNVNVEIDGEGTGA
ENMSIHRAPPTSEBQSNSTELLEITGIKIDLYCPKAGKGVLGAGVGTVMNML
IRNIAIEHSGYSVEFAGSVGERTREGNDFHEKMDENLVDSLVYGOQNEPGRNRA
LTLTMAKFRDEGRDVLFPVDNTIRYTLAETESALLGRMPSAAGVQPTLAEKGL
OERISTKGTSTSVQAVYPADLDTPDSPTTAEHSLGIVSLGILTPAVDP
LESTSRQLDPLVVEGHNVAARGVOTTLQRYKELKIDIIAIGDELSEEDLVYAR
KIERLSQPFVAVFENGTGPKYVPLKETITNGFKGILDEYDHIPEQAFYAGRTIDEV
LEKAKKL"
9436..9864
/gene="atpc"
/note="PM1495"
9436..9864
/gene="atpc"
/codon_start=1
/transl_table=1
/product="atpc"
/protein_id="AAK03579.1"
/db_xref="GI:12721879"
/translation="MSVFNLYVSAEQOIFSGOVESIQATGIEGELGLAGHTPLTA
IKPGIVKLTLEDGKEEVIYVSGGLEVPQNVVTVLADPAIRGDELDGRIILAAKRAE
ENIRRHGDANYEMLASKLSKELAKIRAYELTELKVNKR"
complement(9922..10161)
/gene="PM1496"
complement(9922..10161)
CDS
Query Match 97.6%; Score 1925.2; DB 1; Length 10229;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1941; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 24 gcttgagattctatgatc-acacgaagcaattatcacatggtcaaaagtaattga 82
DB 6794 GCTTCGACTCTTGATGATCAACACGTAAGCAATTAACAAGTCAAAAGTAACGTGA 6853
QY 83 ttattgaacaaacccaactctccgtatctgtagcaacaaagcattagtattt 142
DB 6854 TTATTGAACAAAACAAATACCTCCGTTATCTGTAGCAACAAAGCATTAGCTATT 6913
QY 143 gaagtagagtttgtagtacttagaagaagtgcattagatcgatttgctattgaatca 202
DB 6914 GCAGTAGAGTTGGTACTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6973
QY 203 gcactttgagtagtgcatacattatgctgattttatgctgtagtgaatcccaatct 262
DB 6974 GCACCTTTAGAGTAGCTACCACTAATCTATGCTGATTATGCGTAGTTAACCCCAT 7033
QY 263 ggaattacaatgcataatgaagagtcattaaagagcatttgatagcttcaagca 322
DB 7034 GGCATTAATGATGATCAATTAAGAGTCAATTAAGAGCATTGGATAGCTTCAAGACA 7093
QY 323 aacagtcgttgtagtgcatacatttaattgaagagacaaatgagagtgcttaagaga 382
DB 7094 AACAGTGCCTGTAAGTAACTTAATGAGAGACAAATGCGAGCGCTAAAGAGA 7153
QY 383 taagaacccaatcgcgagtgtaaaagagtcataaagagcatttgatagcttcaagca 442
DB 7154 TAAAGAACCAAAATCGGAGTGTAAAAAAGATACAAAAAATCTTAAAGGAGGAGAAATG 7213
QY 443 ttgcgtcccgaaatgctgaacagcaagagcagtccttccttaagcccttattcag 502
DB 7214 TTGCTGCCCGAAATGCGTAAACGCAAGACGATGCTTCTTACGCCCTTATTCAG 7273
QY 503 aaacaatacgtatgattagccaagttccaagcaagcagaattggttcaagcattcat 562
DB 7274 AAACAATAGCTAGCGATGTAGCCAGCTTCCAAAGCAACGATGTGTTCAAGACATCAT 7333
QY 563 tttagtgatcgcgaagtaaaaaagtgagcatgatttctgttccaagtcgtgctc 622
DB 7334 TTTTAGTGATCGCGAAGTAAAAAGTGGCATGATTTGTGTGCCACAGATCGTGGTC 7393
QY 623 ttgtgtgagcttaaacggaactggtttaaacctgattatgaatgaagaagaatgga 682
DB 7394 TTTGTGTGCTTAAACGTAACCTGTTTAAACCTGTATTAAATGAAATGAAGATGGA 7453

QY	1763	ctctgtcttgcgggggagtgagtgagcgtgaagcgtgaagcgttcatcatgaagt	1822
Db	8534	ctctgtcttgcgggggagtgagtgagcgtgaagcgttcatcatgaagt	8593
QY	1823	gaaagaccttaacgtatagaataagcgtctctgtttatagtcacaaatgaacagacc	1882
Db	8594	gaaagaccttaacgtatagaataagcgtctctgtttatagtcacaaatgaacagacc	8653
QY	1883	aggtaaccgttaccgtctgagcattacaagcgttaactatgcggaataatccgtatga	1942
Db	8654	aggtaaccgttaccgtctgagcattacaagcgttaactatgcggaataatccgtatga	8713
QY	1943	aggctgtagtcttcatctcttcgttgataa	1972
Db	8714	aggctgtagtcttcatctcttcgttgataa	8743
RESULT	2		
LOCUS	U32730/c	13828 bp	DNA BCT 29-MAY-1998
DEFINITION	Haemophilus influenzae Rd section 45 of 163 of the complete genome		
ACCESSION	U32730	142023	
VERSION	U32730.1	GI:3212191	
KEYWORDS	Haemophilus influenzae Rd.		
SOURCE	Haemophilus influenzae Rd		
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Haemophilus.		
REFERENCE	1 (bases 1 to 13828)		
AUTHORS	Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J., Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G.G., Fitzhugh, W., Fields, C.A., Gocayne, J.D., Scott, J.D., Shirley, R., Liu, L.I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T., Hanna, M.C., Nguyen, D.T., Saudke, D.M., Barton, R.C., Fine, L.D., Fritchman, J.L., Fuhmann, J.L., Georgagen, N.S., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O. and Venter, J.C.		
TITLE	Whole-genome random sequencing and assembly of Haemophilus influenzae Rd		
JOURNAL	Science 269 (5223), 496-512 (1995)		
MEDLINE	95350630		
REFERENCE	2 (bases 1 to 13828)		
AUTHORS	Tatusov, R.L., Mushegjan, A.R., Bork, P., Brown, N.P., Hayes, W.S., Borodovsky, M., Rudd, K.E. and Koonin, E.V.		
TITLE	Metabolism and evolution of Haemophilus influenzae deduced from a whole-genome comparison with Escherichia coli		
JOURNAL	Curr. Biol. 6 (3), 279-291 (1996)		
MEDLINE	96396784		
REFERENCE	3 (bases 1 to 13828)		
AUTHORS	White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-JUN-1995) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA		
AUTHORS	White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.		
TITLE	4 (bases 1 to 13828)		
JOURNAL	Submitted (27-SEP-1997) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA		
REMARK	The H. influenzae sequence has been updated by R. Fleischmann. New database matches have been assigned, product names have been improved, and a number of frame shifts have been corrected. We gratefully acknowledge the work of Tatusov et. al. We have incorporated their annotation into the /notes fields of the corresponding H. influenzae genes		
REFERENCE	5 (bases 1 to 13828)		
AUTHORS	White, O., Clayton, R.A., Kerlavage, A.R., Fleischmann, R.D., Peterson, J., Hickey, E., Dodson, R. and Gwin, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-MAY-1998) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA		
REMARK	The whole genome was shifted by 568 nucleotides for a new stat On Jun 12, 1998 this sequence version replaced g1:11573448.		

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 14:58:41 ; Search time 2621.89 Seconds

(without alignments)
11633.746 Million cell updates/sec

Title: US-09-545-199C-3

Perfect score: 1972

Sequence: 1 agcgggccattgctcgaft.....gtctattctctgtgataa 1972

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapept 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba1:*

2: gb_ba2:*

3: gb_ba3:*

4: gb_in1:*

5: gb_in2:*

6: gb_in3:*

7: gb_om:*

8: gb_ov:*

9: gb_pat1:*

10: gb_pat2:*

11: gb_ph:*

12: gb_pl1:*

13: gb_pl2:*

14: gb_pl3:*

15: gb_pl4:*

16: em_ba1:*

17: em_ba2:*

18: em_fun:*

19: em_htgo_hum:*

20: em_htgo_inv:*

21: em_htgo_rod:*

22: em_htg_hum1:*

23: em_htg_hum2:*

24: em_htg_hum3:*

25: em_htg_hum4:*

26: em_htg_hum5:*

27: em_htg_hum6:*

28: em_htg_hum7:*

29: em_htg_hum8:*

30: em_htg_inv1:*

31: em_htg_inv2:*

32: em_htg_other:*

33: em_htg_rod:*

34: em_hum1:*

35: em_hum2:*

36: em_hum3:*

37: em_hum4:*

38: em_hum5:*

39: em_hum6:*

40: em_hum7:*

41: em_in:*

42: em_om:*

43: em_or:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1925.2	97.6	10229	1	AE006185 Pasteurel
2	1255.6	63.7	13828	3	U32730 Haemophilus
3	870	44.1	870	2	AF237922 Pasteurel
4	812	41.2	7141	2	ECOHATP M2546 E. coli H+ A
5	812	41.2	11414	1	AE000450 Escherich
6	812	41.2	14526	2	ECUNC X01631 E. coli ori
7	812	41.2	136254	2	ECOWB2 L10328 E. coli; th
8	810.4	41.1	11999	1	AE005605 Escherich

David L. Lowery
Tony E. Guller
Michael J. Kennedy
April 6, 2002

1184 gttatcaaatgaattaaatgaattcttgcgcgtgcagcagcaatttaacaaatagagg 1243
|||||
7955 gttattcaaatgaattaaatgaattcttgcgcgtgcagcagcaatttaacaaatagagg 8014
|||||
1244 atcgtgaatgcgaactcgtgaaanaatctgacaatcatcgtgtgcgttaattgaactgaatt 1303
|||||
8015 atcgtgaatgcgaactcgtgaaanaatctgacaatcatcgtgtgcgttaattgaactgaatt 8074
|||||
1304 ccccaagaatgcgcgttcaccaaatatataatgcttccttaaatggtgaacacaggttaagtrac 1363
|||||
8075 ccccaagaatgcgcgttcaccaaatatataatgcttccttaaatggtgaacacaggttaagtrac 8134
|||||
1364 tgaagtccaacacaaatgaattggt 1423
|||||
8135 tgaagtccaacacaaatgaattggt 8194
|||||
1424 tgaattaaaacgcgtgttaagcgttaacaaatgaacaaatcccaattctgttccagtggg 1483
|||||
8195 tgaattaaaacgcgtgttaagcgttaacaaatgaacaaatcccaattctgttccagtggg 8254
|||||
1484 aacgaaacacgt 1543
|||||
8255 aacgaaacacgt 8314
|||||
1544 aatcgtgtgcgaagaagaattggt 1603
|||||
8315 aatcgtgtgcgaagaagaattggt 8374
|||||
1604 taacagtaactgaactctttagaagaacgcggaattaaagtatagacttgaattgtgtgtgtgtgt 1663
|||||
8375 taacagtaactgaactctttagaagaacgcggaattaaagtatagacttgaattgtgtgtgtgtgt 8434
|||||
1664 gaaagggggttaagtaagttatcttcgt 1723
|||||
8435 gaaagggggttaagtaagttatcttcgt 8494
|||||
1724 ggaattaatccgttaacatcgcgaattgagcactcagtgtaactctgtcttgcgggggttagg 1783
|||||
8495 ggaattaatccgttaacatcgcgaattgagcactcagtgtaactctgtcttgcgggggttagg 8554
|||||
1784 tgaagcgtgaacgtgtgaagtaacgaactcttcatcagagatgaagaacttaacgaactga 1843
|||||
8555 tgaagcgtgaacgtgtgaagtaacgaactcttcatcagagatgaagaacttaacgaactga 8614
|||||
1844 taaagtgtctctgttcttgaatgaatgaacgaacgcacacgaatgaacgttcttgcgtgtgtgt 1903
|||||
8615 taaagtgtctctgttcttgaatgaatgaacgaacgcacacgaatgaacgttcttgcgtgtgtgt 8674
|||||
1904 attaacaggttaactatgctgcggaanaatctcgtgtgaagaagtcgtgtgtgtgtgtgtgtgtgt 1963
|||||
8675 attaacaggttaactatgctgcggaanaatctcgtgtgaagaagtcgtgtgtgtgtgtgtgtgtgt 8734
|||||
1964 cgtgtgataa 1972
|||||
8735 cgtgtgataa 8743
|||||

RESULT 2
AF237922 870 bp DNA BCF 08-MAY-2000
LOCUS Pasteurella multocida ATP synthase F1 gamma chain (atpg) gene,
DEFINITION complete cds.
ACCESSION AF237922
VERSION AF237922.1 GI:7716504
KEYWORDS
SOURCE
ORGANISM
Pasteurella multocida.
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Pasteurella multocida
Pasteurella
1 (bases 1 to 870)
REFERENCE
AUTHORS
TITLE
Fuller,T.E., Kennedy,M.J. and Lowery,D.E.
Identification of Pasteurella multocida virulence genes in a
septicemic mouse model using signature-tagged mutagenesis

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 870)
AUTHORS Fuller,T.E., Kennedy,M.J. and Lowery,D.E.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-2000) Discovery Research, Pharmacia & Upjohn
Animal Health, 7923-25-434, 7000 Portage Road, Kalamazoo, MI
49001-0199, USA
FEATURES
source
1..870
/organism="Pasteurella multocida"
/db_xref="taxon:747"
1..870
gene
1..870
/gene="atpg"
CDS
1..870
/gene="atpg"
/codon_start=1
/transl_table=11
/product="ATP synthase F1 gamma chain"
/protein_id="AF68408.1"
/db_xref="GI:7716505"
/translation="MAGKEIRKIASVSTOKITKAMEVVAASKRKQERMSRRP
YSEIINNVISHSKATIGKHPVLDREVKVMIVSTDRGLGGLNLPKTYNE
MKEMKEKDIVSVOLSLIGSKSINFPOLGILILQDSIGIDTDPSEDLISVMSIDAY
KKEGVVYLVYVKFINTMSOKVLEKLIPLDELNDDELGERQVMDVYEPDAKVL
DNILVRYLESQVYQAAVENLASDAQARVMYAMKATDNAGNLINELQLVYNAKQASIT
NELNEVAGCAAT"
BASE COUNT 305 a 132 c 177 g 256 t
ORIGIN
Query Match 44.1%; Score 870; DB 2; Length 870;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 364 atgacagtgcttaagaagataagaacccaanaatcgcgagtgtaaaagttacacaaaatt 423
|||||
DB 1 ATGGCAGGTGCTTAAGAAGATPAAGAACCAAAATCCGAGTGTAAAGTACCAAAAAATT 60
QY 424 actaaagcgtatgaatggt 483
|||||
DB 61 ACTAAAGCGATGGAATGCTGTGCTCGAATAATGCGTAAAGCGAAGACCATGTCT 120
QY 484 tcttcacgccccttcttgaagaacatacgtacgt 543
|||||
DB 121 TCTTACGCCCTTATTTCGAAGACATACGTACCGATGATTACCAAGTTCACAAAGCAACG 180
QY 544 atggttacaagcatccattttagtgaatcgcgaagtaaaaaagttggtgatgtt 603
|||||
DB 181 ATGTGTACAAGCATCCATTTTATGTAGTGCAGCAATGAAAAAAAGTGGCATGATTGTT 240
QY 604 gtgtccacagatcgt 663
|||||
DB 241 GTGTCCACAGATCGGTGCTTTGTGTGTGTAAACGTAACTGTTTAAACGTATTA 300
QY 664 aatgaatgaagaatgaagaagaagaatgttccgttcaattgattatcgtgtct 723
|||||
DB 301 AATGAATGAAGAATGGAAGAAAAAAGATGTTCCGTTCAATTGATGTTAATCGTGTCT 360
QY 724 aaatcatcaacttttccaatcttgggaattaaatttaacccaagatcaagttct 783
|||||
DB 361 AAATCTATCAACTTTTCCAACTTTGGGAATTAAATTTAAACCAAGATTCAGGTATT 420
QY 784 ggtgtactccctcgtgtggaagtaattgtgtgaatcgaatcttgaattgaattgtctt 843
|||||
DB 421 GGTGATACCTCCCTCGTGTGAGCAGTTAATGTTAGTCATTTAGTTGATCTTAT 480
QY 844 aaaaaaggaagtagatgt 903
|||||
DB 481 AAAAAAGGGAAGTAGATGTTGTGATTTAGTTTAAACAAATTTATTAACGATGTCG 540
QY 904 caaaagcagtatgtgaaanaatattccattcagaagattagataatgataatgagc 963
|||||
DB 541 CAAAGCCAGTGTGGAATAATTAATTCATTACAGATTAATGATGATGATTAAGGC 600
|||||

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 15:36:06 ; Search time 2621.23 Seconds

(without alignments)
11636.675 Million cell updates/sec

Title: US-09-545-199c-3
Perfect score: 1972
Sequence: 1 agcgggcattgctcagf.....gtctattcttcgtgataa 1972

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1344157 seqs, 7733874588 residues

Word size : 25

Total number of hits satisfying chosen parameters: 12

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rtd:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rtd:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_rtd:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vl:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vl1:*
59: gb_vl2:*
60: gb_hcg1:*
61: gb_hcg2:*
62: gb_hcg3:*
63: gb_hcg4:*
64: gb_hcg5:*
65: gb_hcg6:*
66: gb_hcg7:*
67: gb_hcg8:*
68: gb_hcg9:*
69: gb_hcg10:*
70: gb_hcg11:*
71: gb_hcg12:*
72: gb_hcg13:*
73: gb_hcg14:*
74: gb_hcg15:*
75: gb_hcg16:*
76: gb_hcg17:*
77: gb_hcg18:*
78: gb_hcg19:*
79: gb_hcg20:*
80: gb_hcg21:*
81: gb_hcg22:*
82: gb_hcg23:*
83: gb_hcg24:*
84: gb_hcg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_rtd1:*
95: gb_rtd2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1674	84.9	10229	1	AE006185 Pasteurel
2	870	44.1	870	2	AF237922 Pasteurel
3	44	2.2	1374	9	AR089428 Sequence
4	44	2.2	13828	3	U32730 Haemophilus
5	35	1.8	307	1	AF032869 Actinobac
6	32	1.6	793	1	AF037146 Salmonell
7	32	1.6	793	1	AF037147 Salmonell
8	32	1.6	793	1	AF037148 Salmonell

9 32 1.6 793 1 AF037149
10 32 1.6 793 1 AF037150
11 26 1.3 7307 2 BACATPA M20255 B.megateriu
c 12 25 1.3 34699 1 AF008210 AF008210 Buchnera

ALIGNMENTS

RESULT 1
AE006185 10229 bp DNA BCT 08-MAR-2001
LOCUS Pasteurella multocida pm70 section 152 of 204 of the complete
DEFINITION genome.
ACCESSION AE006185 AE004439
VERSION AE006185.1 GI:12721868
KEYWORDS
SOURCE Pasteurella multocida.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Pasteurella.
REFERENCE 1 (bases 1 to 10229)
AUTHORS May,B.J., Zhang,Q., Li,L.L., Paustian,M.L., Whittam,T.S. and
Kaput,V.
TITLE Complete genomic sequence of Pasteurella multocida, pm70
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)
MEDLINE 21145866
REFERENCE 2 (bases 1 to 10229)
AUTHORS Zhang,Q. and Kaput,V.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2000) Department of Veterinary Pathobiology,
University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN
55108, USA

FEATURES

source location/Qualifiers

1..10229 /organism="Pasteurella multocida"

/strain="PM70"

/db_xref="taxon:747"

275..2164

/gene="g1d4"

/note="PM1485"

275..2164

/gene="g1d4"

/codon_start=1

/transl_table=1

/product="g1d4"

/protein_id="AAK03569.1"

/db_xref="GI:12721869"

/translation="MFTYENVDIVIVIGGHANGTEALAPARMGLKTLILHNVDTLQ
MSCPAIGIGIGKHLVREIDAMGGLMATADQAGIQPTLNSSKGPVAVRATRADRV
LYQPAVIALENGENLIDFOEVDIILDQDRVGVYTKMGLFKHAKAVILLTACTFIS
GKHIGLENTGGRAGDPASVAMLDRLRELRLVDRLKTGPRIIDARTIDFVLAQ
HGDKELVYFSPMGSDVQHPRIQCFITHTNEQRLRYNNLDRSPMAGIIEGIPRY
CPSTEDKVMRSEKNSHQITLDEPGLSNETIPNGISTSLPFDYQAKTVNSKMEKA
RIKPGVAIEVDYDPDLKPTLETKSIRGLFFAGQNGTGYEAGOGLLAING
LFVQEKAMPFRDQAYIGVLYDLCITGKEPRVYTSRAEYRLRLRENDADRLLP
IAHQIGLDEKRAKRFNOKMENIELEORRLMLHPREYLDANKVILGSPILVRES
GEDLRRPEMYOITLSLTPROPAMDQEAQVEYIAIKQGYIEHOOEELAROKRHE
STALPAHFDTVVSGLSNEVRAKLEQHRPVSIGASRLSGVTTPAISTILVLSKKQGM
LKRGE"

2164..2796

/gene="g1d4"

/note="PM1486"

2164..2796

/gene="g1d4"

/codon_start=1

/transl_table=1

/product="g1d4"

/protein_id="AAK03570.1"

/db_xref="GI:12721870"

/translation="MTNDEQLSQKLETLKQTAISTTDQKEQIKVILLNKRNA
YNLTSVRDPMELVKHILDSVYSPHLQGHFIDVGTGPGIPGLPLAVNPNKHFVLL
DSLGRKISFIRNAIRELGIDNVEAVLSHVEEYHPEQPPDGLVSRASFSLKDMTDCOH

gene

CDS

gene 2934..3341
/gene="PM1487"
2934..3341
/gene="PM1487"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAK03571.1"
/db_xref="GI:12721871"
/translation="MSKVYQAKKQKKTALGIELOVLASLICILGYQOVAISPLIG
SSAIFPHCLFVWFVEFTKKQKSNKLTTFYGGVAVKISTLFLIVLAKLFNSMFI
AFVGVFTIVINLILPFISKYLNIKHFHYQ"
3352..4146
/gene="atpb"
/note="PM1488"
3352..4146
/gene="atpb"
/codon_start=1
/transl_table=1
/product="atpb"
/protein_id="AAK03572.1"
/db_xref="GI:12721872"
/translation="MAELTTAGYICGHLALFKTGDSFWHVLDTLLFSIIGAIPLF
VSKVAKKAPPGVSKMOCEVEIMVDIGIYKENGGRHAGVPLATIFGWFIMN
AIDLIVDDLPLQALHFGIEYLRVPTADISGLTIGLISGIFLIFYTKSGMSGFV
KETTLPENHPLIPVNLALLESYTLAKRVSLAFRFGNMYGSELFIILAVMYMANN
FALNSMGIFMHLMAIFHLIVTLQAFIFMMLTVYLSKGYAKAEH"
4203..4457
/gene="atpb"
/note="PM1489"
4203..4457
/gene="atpb"
/codon_start=1
/transl_table=1
/product="atpb"
/protein_id="AAK03573.1"
/db_xref="GI:12721873"
/translation="MEVITPTTIASLILAVNALGTALGFAILGKFLSSARQPEL
ASSLIQIMFTVAGLDAISMVAVGIALLPFIANPFDLKL"
4537..5007
/gene="atpf"
/note="PM1490"
4537..5007
/gene="atpf"
/codon_start=1
/transl_table=1
/product="atpf"
/protein_id="AAK03574.1"
/db_xref="GI:12721874"
/translation="MNINATLIGLILAFALFVAFCKMFWMPPIIKATEEORSIANTL
ASMAAKKEQADVTALVEQETTEKMQAQIIDLANKRRNILEEYVEAEARKKATI
EOGAEVAEARKRQVEELRVKVASLAIAGAEKIVGRVDRANSDIIDKLVAEL"
5021..5554
/gene="atph"
/note="PM1491"
5021..5554
/gene="atph"
/codon_start=1
/transl_table=1
/product="atph"
/protein_id="AAK03575.1"
/db_xref="GI:12721875"
/translation="MSELTITAPRYAKAFDFAIQNAVEKMANMLHFSSELLIKDETLM
OTFLKSSLSAKLADVTYISCGEQLDYGONLIRLMAENKRLVLPALPIYQAFQHYVE
HOKYAEVOYSAOPLNATQEQKIASAEKRLKRVKINCISLDSIAGVIIIRDDFVI
DGSSRGOLSRANLQJL"
5567..7108
/gene="atpa"
/note="PM1492"
5567..7108
/gene="atpa"
/codon_start=1

gene

CDS

gene

CDS

gene

CDS

gene

CDS

OY 964 gaagaagaacagtttgagattatattacgaaccgtatgcgaagattattatgaatt 1023
 DB 601 GAAAGAAAAACAGTTGGATTATATTATTCGAAACCGATGCGAAATATTATTAGTAAT 660
 OY 1024 ttatgtgtcgttattatagaatcaggtttatcaagcagcaggttgaacattgcttct 1083
 DB 661 TTATGTGTTGCTTATTATTAATCTCAGGTTTATCAAGCAGCAGTTGAAACCTGCTTCT 720
 OY 1084 gagcgaagcgcgtcgaatggtgcgcattgaagaagcagaataagcaggttaatt 1143
 DB 721 GAGCAAGCGCGCATGATGCTCCGATGAAGCAGCAACGATACGCACTAAT 780
 OY 1144 aatgagttacagttatgctatatacaagaagctgtaagcagattatcaatgaatt 1203
 DB 781 AATGAGTTACAGTTATGCTATATCAAAAGCTGCTCAAGCAAGTATTAATGATTAAT 840
 OY 1204 gaattgttcgcgtgcagcagcaatttaa 1233
 DB 841 GAAATTTGTCGCGTGACGACCAATTTAA 870
 RESULT 3
 LOCUS AR089428 1374 bp DNA PAT 07-SEP-2000
 DEFINITION Sequence 187 from patent US 5994066.
 ACCESSION AR089428
 VERSION AR089428.1 GI:10016185
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1374)
 AUTHORS Bergeron,M.G., Picard,F.J., Ouellette,M. and Roy,P.H.
 TITLE Species-specific and universal DNA probes and amplification primers to rapidly detect and identify common bacterial pathogens and associated antibiotic resistance genes from clinical specimens for routine diagnosis in microbiology laboratories
 JOURNAL Patent: US 5994066-A 187 30-NOV-1999;
 FEATURES
 SOURCE location/Qualifiers
 1. 1374
 BASE COUNT 425 a 238 c 312 g 399 t
 ORIGIN
 Query Match 2.2%. Score 44: DB 9; Length 1374;
 Best Local Similarity 100.0%; Pred. No. 1.5e-12;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1857 gttatgtcaaatgaacgagccagcaggttaccgttaccgtt 1900
 DB 607 GTTATGTGTCAAATGACGACGACCAAGTACCGTTTACGTGT 650
 RESULT 4
 LOCUS U32730 13828 bp DNA BCT 29-MAY-1998
 DEFINITION Haemophilus influenzae Rd section 45 of 163 of the complete genome.
 ACCESSION U32730 L42023
 VERSION U32730.1 GI:3212191
 KEYWORDS
 SOURCE Haemophilus influenzae Rd.
 ORGANISM Haemophilus influenzae Rd.
 Bacteria; Proteobacteria; gamma subdivision: Pasteurellaceae;
 Haemophilus.
 1 (bases 1 to 13828)
 Fleischmann,R.D., Adams,M.D., White,O., Clayton,R.A.,
 Merrick,J.M., McKenney,K., Sutton,G.G., FitzHugh,W., Fields,C.A.,
 Gocayne,J.D., Scott,J.D., Shirley,R., Liu,L.I., Glodek,A.,
 Kelley,J.M., Weidman,J.F., Phillips,C.A., Spriggs,T., Hedblom,E.,
 Cotton,M.D., Utterback,T., Hanna,M.C., Nguyen,D.T., Saudek,D.M.,

Haemophilus influenzae

Brandon,R.C., Fine,L.D., Fritchman,J.L., Fuhrmann,J.L.,
 Geisler,N.S., Gnehm,C.L., McDonald,L.A., Small,K.V., Fraser,C.M.,
 Smith,H.O. and Venter,J.C.
 Whole-genome random sequencing and assembly of Haemophilus
 influenzae Rd
 Science 269 (5223), 496-512 (1995)
 JOURNAL MEDLINE 9530630
 REFERENCE 2 (bases 1 to 13828)
 Tatusov,R.L., Mushegian,A.R., Bork,P., Brown,N.P., Hayes,W.S.,
 Borodovsky,M., Rudd,K.E. and Koonin,E.V.
 Metabolism and evolution of Haemophilus influenzae deduced from a
 whole-genome comparison with Escherichia coli
 Curr. Biol. 6 (3), 279-291 (1996)
 JOURNAL MEDLINE 96398784
 REFERENCE 3 (bases 1 to 13828)
 White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.
 Direct Submission
 Submitted (25-JUL-1995) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 JOURNAL MEDLINE 9712
 REFERENCE 4 (bases 1 to 13828)
 White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.
 Direct Submission
 Submitted (27-SEP-1997) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 The H. influenzae sequence has been updated by R. Fleischmann. New
 database matches have been assigned, product names have been
 improved, and a number of frame shifts have been corrected. We
 gratefully acknowledge the work of Tatusov et. al. We have
 incorporated their annotation into the /notes fields of the
 corresponding H. influenzae genes
 5 (bases 1 to 13828)
 White,O., Clayton,R.A., Kerlavage,A.R., Fleischmann,R.D.,
 Peterson,J., Hickey,E., Dodson,R. and Gattin,M.
 Direct Submission
 Submitted (28-MAY-1998) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 The whole genome was shifted by 588 nucleotides for a new start
 On Jun 12, 1998 this sequence version replaced g1:1573448.
 COMMENT Location/Qualifiers
 FEATURES
 SOURCE
 1. 13828
 /organism="Haemophilus influenzae Rd"
 /db_xref="taxon:71421"
 65. 1168
 /gene="HI0470"
 65. 1168
 /note="Similar to SP:P06986 GB:U02071 GB:X03416 PID:41695
 PID:41710 percent identity: 60.40; identified by sequence
 similarity: putative"
 /codon_start=1
 /transl_table=11
 /product="histidinol-phosphate aminotransferase (hisC)"
 /protein_id="AAC22129.1"
 /db_xref="GI:1573449"
 /translation="MNDYNSHILQWKRTITITLSRONIOALPPYOSARKLGNGTI
 MLNNEYTPSPFOLSGKDLNRYPEPOPOVOAYANAGSTENVLVTREGDEIEL
 IITPTEPKDAIIECPPTCYGMYASATACVLSVTYLPDPDQINLPEIKHNLDYK
 VVFCSPNNPFGNLEKQSDILDLQITGKATVYDEAVTECEPASVINILKNTPL
 AITFLSKAFALAGRCGVLANPELIDILSKVIAPYIPVPSADLAQALRPSIAT
 VOALTOELISNRQWIAKALIVLHVEKEYESANVLLIKONGQAVFRAWEQGIIR
 DONKTLHONCIRITVGTGRNCEKVEAIKFEVK"
 1280. 2368
 /gene="HI0471"
 1280. 2368
 /note="Similar to SP:P06987 GB:X03416 PID:41696 PID:41711
 GB:U00096 percent identity: 64.97; identified by sequence
 similarity: putative"
 /codon_start=1
 /transl_table=11
 /product="imidazoleglycerol-phosphate dehydratase /
 histidinol-phosphatase (hisB)"
 /protein_id="AAC22130.1"

/db_xref="GI:1573450"
 /translation="MQPLFLIDRGDTLIDEPKTDFOIDSLEKLEKLEPKYIPALLYLKA
 KREIVASMODGLTDAFQPDGPHNYMALFESOGITFDEVLICPKRPECNLCR
 KPKILDIHYIKRNLEEDIRSFVIGDRETDVLAENGLIRAIQYDPOKMMDLIAEKL
 LGEYTNLGGKRRPRAEVIROKTEIDYQVWLDEAGVEITGVGFEDHMDLIATLH
 GGFRRNVQCKGDLMDENHTVDTALALQALQAKQAVGDRGIARQGFVLPMDECAEC
 ALDLSGRPIKFNACFKRDVGFSTELTEHFQSLAFSLMTLHLNVTGNNDHRIE
 SLFKAFGRTRQAIRIEGNEMPSKGVLT"
 2434..3033
 /gene="HI0472"
 2434..3033
 /gene="HI0472"
 /note="similar to SP:P10375 PID:41712 GB:U00096
 PID:1736701 PID:1736710 percent identity: 55.90;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="amidotransferase (hisH)"
 /protein_id="AAC22131.1"
 /db_xref="GI:1573451"
 /translation="MINITIIDTCANLSSVKAFAFDRLGNTETEDLNKIKSADKLI
 LKCVGTAAMAYNLQEROLIETIONTLPVLGICIGMOTERSEGNVPTNLISGK
 TRNIPDTGLPLPMGNRRVQFYKNCPLFDGIYONSHFVHSVSPNHSVAISVYG
 VNESAIAKENFYGVQFHPERSGKNGALLKFNKVPF"
 3069..3818
 /gene="HI0473"
 3069..3818
 /gene="HI0473"
 /note="similar to SP:P10371 PID:41713 GB:U00096
 PID:1736702 PID:1736711 percent identity: 62.86;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="phosphoribosylformimino-5-aminoimidazole
 carboxamide ribotide isomerase (hisA)"
 /protein_id="AAC22132.1"
 /db_xref="GI:1573452"
 /translation="MKOSIIIPALDLINGVVRILHOGDYAKQTTYSNDPIKQFNDYVR
 OGKAKHLVDLGCANPOSROFALIGKIVATCGKQVGGIGTEDVDADLLVGNR
 VVIGSTVAHRSVKNMFKIKYGAKEFVLADVNINSGKIVASIGMOEBSGLLETL
 IEPQTVGLQOVCTDISRGTGSGNIGCYDICEKPYPIQSSGIGSLADIEAL
 KGTGSGVIVGRLEBGTFLSEAIKQWQNG"
 3800..4576
 /gene="HI0474"
 3800..4576
 /gene="HI0474"
 /note="similar to SP:P10373 PID:41714 PID:598474 GB:U00096
 PID:1736704 percent identity: 82.03; identified by
 sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="hisF cyclase (hisF)"
 /protein_id="AAC22133.1"
 /db_xref="GI:1573453"
 /translation="MLAKRIIPCLIDVGDGVKGVQFRNHETICDITPLAQRVQESA
 DELVFDITASDSGRKIVADIDIPCVAGIKTTEAEKLFARFAGAKISL
 NSPLADPTLISADRFQVAYVIGDSDFGVWVQYDGTGRTKQRMWOLL
 DMWEYQCGAGETIVLMMNODGLRNGYDIAOLKLVGVCRVPLIASGAGKMWIFRD
 AETAKYDGLASVFFHKIIEIGELKSYLVQSAIEIRSE"
 4576..5241
 /gene="HI0475"
 4576..5241
 /gene="HI0475"
 /note="similar to SP:P06989 GB:D43637 GB:U02072 GB:X03974
 PID:41700 percent identity: 61.22; identified by sequence
 similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="phosphoribosyl-AMP cyclohydrolase /
 phosphoribosyl-ATP pyrophosphohydrolase (hisB)"
 /protein_id="AAC22134.1"
 /db_xref="GI:1573454"
 /translation="MNITKIDMWKVNGLLPVYIIONISTREVLMGYNNEALTTTKE

RKVFPSRTKQRLMTKGEISGNFLNVEEMSLDCNDTLLILVDPIGATCHGEYSCFH
 DEFSQSENKQOAFNMWAMFIKLEQHLKRNKADPSNSYTTALHAKGTRKIAQKVEE
 GVETALAIVADQKAVISEADVLVHLVTVLHNLQDLQWVEITAKLOERHOGIGLHPGEC
 GNM"
 5323..5475
 /gene="HI0476"
 5323..5475
 /gene="HI0476"
 /note="hypothetical protein; identified by Genemark;
 putative"
 /codon_start=1
 /transl_table=11
 /product="H. influenzae predicted coding region HI0476"
 /protein_id="AAC22144.1"
 /db_xref="GI:1573464"
 /translation="MVIKCIDQKQNLNIIPLLLKQOYKREDSKKFTYKRYLQTVN
 YRIQLS"
 5505..6707
 /gene="HI0477"
 5505..6707
 /gene="HI0477"
 /note="similar to GB:M23240 SP:P18199 PID:148089 GB:U00096
 PID:1736569 percent identity: 47.24; identified by
 sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="tyrosine-specific transport protein (tyrP)"
 /protein_id="AAC22135.1"
 /db_xref="GI:1573455"
 /translation="NKKTVGSTLLVAGTIGAGMLAMPLTSAGIGFTVLVLLGLMA
 LTFPSALLEVEIKQTAESDAGIGTIAEQYFGKGRIRIAVAILFPLAIAVYISGG
 SILKDLIPESFGKVSLLFTVIFGSEIYGFHSVDKINRVLEFFVLAFAVLSLML
 PEIKFNLNATPIIDKALLISASPVFTFAGFHSIPSLKKYLDGNVKALREYLSVSA
 ITCAYILMQLSTHGLTQNEFLQILKEBRTNLGLYKATPATGSSVNTASVLTSTL
 ALITSPILGVGLGLEICIEDLKRSFVNTAGRTSLGLTTPILPVLPLFTPBEGTILAG
 YAGOMPAFAVVLPSVLWKKARRAHNLPYKVGWMLTILVVLGLTITSIPFAIRA
 GYLPFVYG"
 complement(6893..7321)
 2.28; Score 44; DB 3; Length 13828;
 Best Local Similarity 100.0%; Pred. No. 1,6e-12;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 gene
 Query Match
 Best Local Similarity 100.0%; Pred. No. 1,6e-12;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1857 gttatggtcaaatgaacgagccacggtacggtttacgtt 1900
 |||||||
 Db 8118 GTTATGTCCTAATGAACGACGACGTAACCGTTACGTGT 8075
 RESULT 5
 AF032869 307 bp DNA BCT 27-NOV-1997
 LOCUS
 DEFINITION Actinobacillus pleuropneumoniae ATP synthase gene, partial cds.
 ACCESSION AF032869
 VERSION AF032869.1 GI:2641642
 KEYWORDS
 SOURCE
 ORGANISM
 Actinobacillus pleuropneumoniae.
 Actinobacillus pleuropneumoniae
 Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 Actinobacillus.
 1 (bases 1 to 307)
 AUTHOR Baumeister,C.X. and Fenwick,B.W.
 TITLE Direct Submission
 JOURNAL
 Submitted (03-NOV-1997) Diagnostic Medicine/Pathobiology, Kansas
 State University, 1800 Denison Ave, Manhattan, KS 66506-5605, USA
 FEATURES
 source
 1..307
 /organism="Actinobacillus pleuropneumoniae"
 /strain="4074"
 /serotype="1"
 /db_xref="taxon:715"
 1..>307
 /codon_start=1
 /transl_table=11

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 14:58:41 ; Search time 74.92 Seconds
(without alignments)
4962.936 Million cell updates/sec

Title: US-09-545-199C-3

Perfect score: 1972

Sequence: 1 agcggcgccattgcgcagf.....gtctattcttcgttataa 1972

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, NA: *

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/6C.COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/6D.COMB.seq:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	555.6	28.2	1374	2	US-08-743-637B-187
2	48.8	2.5	7218	1	US-08-232-463-14
3	41.8	2.1	872	4	US-08-998-416-487
4	40.4	2.0	3095	6	5231168-1
5	40.2	2.0	2107	4	US-08-235-836C-73
6	38.8	2.0	1899	1	US-08-634-331-2
7	38.6	2.0	1991	4	US-08-235-836C-69
8	38.6	2.0	1991	4	US-08-235-836C-77
9	38	1.9	5361	4	US-08-973-462-2
10	38	1.9	6152	4	US-08-973-462-1
11	36.4	1.8	1689	1	US-07-991-867B-41
12	36.4	1.8	1689	2	US-08-544-332-41
13	36.4	1.8	8457	2	US-07-991-867B-1
14	36.4	1.8	8457	2	US-08-544-332-1
15	36.2	1.8	9636	1	US-08-323-170B-1
16	36.2	1.8	10395	1	US-08-245-809-3
17	36.2	1.8	10396	1	US-08-245-809-5
18	36.2	1.8	10798	1	US-08-107-748-2
19	36.2	1.8	10798	5	PCT-US92-01385-2
20	36.2	1.8	10965	1	US-08-107-748-4
21	36.2	1.8	10965	5	PCT-US92-01385-4
22	36	1.8	1152	2	US-08-286-819A-13
23	36	1.8	1152	3	US-08-980-357-13
24	36	1.8	2081	4	US-08-235-836C-71
25	36	1.8	7225	2	US-08-286-819A-15
26	36	1.8	7225	3	US-08-980-357-15
27	36	1.8	10851	2	US-08-286-819A-16

28	36	1.8	10851	3	US-08-980-357-16	Sequence 16, Appl
29	36	1.8	19124	2	US-08-487-826B-13	Sequence 13, Appl
30	35.6	1.8	959	2	US-08-634-331-1	Sequence 1, Appl
31	35.6	1.8	959	6	5508165-1	Patent No. 5508165
32	35.6	1.8	2570	2	US-09-056-075-2	Sequence 2, Appl
33	35.4	1.8	1146	4	US-09-277-716-21	Sequence 21, Appl
34	35.4	1.8	2319	2	US-08-966-389-5	Sequence 5, Appl
35	35.4	1.8	2319	2	US-09-103-509-5	Sequence 5, Appl
36	35.4	1.8	2319	2	US-09-102-644-5	Sequence 5, Appl
37	35.4	1.8	2319	2	US-09-218-032-5	Sequence 8, Appl
38	35.4	1.8	3345	2	US-08-966-389-8	Sequence 8, Appl
39	35.4	1.8	3345	2	US-09-103-509-8	Sequence 8, Appl
40	35.4	1.8	3345	2	US-09-102-644-8	Sequence 8, Appl
41	35.4	1.8	3345	2	US-09-218-032-8	Sequence 8, Appl
42	35.4	1.8	19124	2	US-08-487-826B-13	Sequence 13, Appl
43	34.8	1.8	1035	2	US-08-417-210A-111	Sequence 11, App
44	34.8	1.8	2542	4	US-09-330-317B-9	Sequence 9, Appl
45	34.8	1.8	7101	1	US-08-480-604A-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-08-743-637B-187
Sequence 187, Application US/08743637B
Patent No. 5994066
GENERAL INFORMATION:
APPLICANT: BERGERON, Michel G.
APPLICANT: PICARD, Francois J.
APPLICANT: COELLETTTE, Marc
APPLICANT: ROY, Paul H.
TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA
TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND
TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED
TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS
NUMBER OF SEQUENCES: 273
CORRESPONDENCE ADDRESSES:
ADDRESSEE: QUARLES & BRADY
STREET: 411 EAST WISCONSIN AVENUE
CITY: MILWAUKEE
STATE: WISCONSIN
COUNTRY: USA
ZIP: 53202-4497
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/743,637B
FILING DATE: 04-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/526,840
FILING DATE: 11-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: BAKER, Jean C.
REGISTRATION NUMBER: 35,433
REFERENCE/DOCKET NUMBER: 850586.90012
TELEPHONE: (414) 277-5000
TELEFAX: (414) 277-5591
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
LENGTH: 1374 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Haemophilus Influenzae
US-08-743-637B-187

	Query Match	28.2%	Score 555.6;	DB 2	Length 1374;	
	Best Local Similarity	85.6%;	Pred. No. 3.6e-135;			
	Matches 618;	Conservative	0;	Mismatches 104;	Indels	0; Gaps
OY	1251 atggaactcggaaaaaatgtlacaataatcggtgcgttatattgaacgttgaattcccaaa	1310				
Ddb	1 ATGTCAGACAGGAAAAATGTCAAAATCATCGTGGTGAGTTGCCTGAATTCCCAA	60				
OY	1311 gatcgaatccaaaagtatatagatgcccttaattttaaacagaatttagtacctgaagt	1370				
Ddb	61 GATGAGAGTCCTAAAAGTTTAGCATTAATAAAGTTGATCAGGTTTTAACACTTGAAGTG	120				
OY	1371 caacacaattagtcgtgcgttgatgctgcgtlcatcgcaalvggatacatcbatygatta	1430				
Ddb	121 CACACAACATTATAGTGGGGAGTATTCGTGTTATGTCATTGGGTACATCGATGAGTTTA	180				
OY	1431 aaacgcgyltaaagcgtacaacaatacgaaatccccaatllctgtlccaaglyggaaagaa	1490				
Ddb	181 AAACGGCGCTTTAAAGTGTGAAAACCAATTAACCAATTCAGTACAGTAGAGCACAAA	240				
OY	1491 acattgggtcgtatataatgaacgatattgggtbaaccaatcga tgaacgaagtygaatcggt	1550				
Ddb	241 ACGCTTGGTCGATTTATGCAATGTGTTGGGCCAACCAATTGACGAACAAGAAGCAATTGGT	300				
OY	1551 gcagaaagaatlygtlctatlaccgtygcaccaagaattatgaaagaaatctaacaagt	1610				
Ddb	301 GAAGAGAGACGTTGGGCTATCCACCCTTCGGCACCCAGCTATGAAGAACATCAAAACAGT	360				
OY	1611 acggaactttgaaacgggaattaaagtatacgatctagtttgcgtlttcggaagagg	1670				
Ddb	361 ACGGAATTTATTGAACCTGTTATCAAAAGATCGATCATTTATTTGCCATTTGCAAAAAGGT	420				
OY	1671 ggtaaagtatgattatcgcgtgtygcggtylgtcgglaaaacgcgtcaatatagatygaaat	1730				
Ddb	421 GGTAAGTGTGTTATTATTCGGCGGTGCAGGCTGTAGGTAAACCGTAAATATATGGAGTTG	480				
OY	1731 atccgtaacatcgcgaattagcacctcaagttactctgtctttgsggggtatgagtcgt	1790				
Ddb	481 ATTGCTAATATTCGATTGACATTTAGGTATCTGTGTTTGGCGGAGTACGCAACAGT	540				
OY	1791 acgcgtgaagtacaactcttcatcatgatagtaaagaaactcaacgtatagataaagt	1850				
Ddb	541 ACCTGCTGAAGCTAATACATCTATCATGAAATGAAGAATTTAAAGTATAGATAAAGTA	600				
OY	1851 tctcttgttataatgtaaatgaagccacgaagtaacgcgtttacgtgtgtggtcataa	1910				
Ddb	601 TCTTGGCTTTATGCTAATAATGAACGACGACCAAGATMAACCGTTTACGCTGTGCATTTGACT	660				
OY	1911 ggcttaactatgcccgaaaaatctcgttgaagaagtcgtgatacttattcttcgttat	1970				
Ddb	661 GGTTAACATGCTGCTGAAAATTTCCGTGATGAAGTCGAGATGTATTTCTTTTGGAT	720				
OY	1971 aa 1972					
Ddb	721 AA 722					
	RESULT 2					
	US-08-232-463-14/c					
	; Sequence 14, Application US/08232463					
	; Patent No. 5670367					
	GENERAL INFORMATION:					
	APPLICANT: DORNER, F.					
	APPLICANT: SCHEIFLINER, F.					
	APPLICANT: FALKNER, F. G.					
	TITLE OF INVENTION: RECOMBINANT FOMLOPX VIRUS					
	NUMBER OF SEQUENCES: 52					
	CORRESPONDENCE ADDRESS:					
	ADDRESSEE: Foley & Lardner					
	STREET: 1800 Diagonal Road, Suite 500					
	CITY: Alexandria					

STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PT29pt-F15

[illegible]

RESULT 5
 US-08-235-836C-73
 ; Sequence 73, Application US/08235836C
 ; Patent No. 6248562
 ; GENERAL INFORMATION:
 ; APPLICANT: Dunn, John J.
 ; APPLICANT: Luft, Benjamin J.
 ; TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
 ; TITLE OF INVENTION: Borrella Polypeptides and Uses Therefor
 ; NUMBER OF SEQUENCES: 144
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Brookhaven National Laboratory
 ; STREET:
 ; CITY: Upton
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 11973
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/235,836C
 ; FILING DATE: 29-Apr-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/148,191

US-08-235-836C-69
Sequence 69, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Poly peptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235.836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogoslian, Margaret C.
REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BN/93-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729

US-08-235-836C-77

Query Match	2.0%	Score	38.6	DB	4	Length	1991
Best Local Similarity	48.1%	Pred. NC	0.76				
Matches	139	Conservative	0	Mismatches	149	Indels	1
						Gaps	1

OY	1110	gaagcagcaacagatatacgcaagtacttcaatatgagttacagtttagtcataaca	116
Db	654	GAAATGAATCAGGTTTAACTTTTGCAAGAGATTACAGCAATTCAGCGAATCTATA	713
OY	1170	agctgcgaagcaaglatltaacaatgaattaaatgaattgttcggcgltcacgaacat	1222
Db	714	AGCGAGATCAAGATTAATTTGATATGTAGATTAGATTAATTTTCATGAAGAAGTATCCAAATAT	773
OY	1230	ttacaacaatgagatctcggtcaatggcgcaactcggaanaatttacaacatcatcggtgc- ggt	128
Db	774	AACGGAAACTATTTGACGATTTTAAAGGGATCGCTTGAAAANAAGTACAGATGAAGCATPA	833
OY	1289	tattgacgttgaattccccaagaatgcagfaccaaaaatlatalgatgccttlaaatgtlga	134
Db	834	AAAAAGATTTGAAAGTCAGCGTGTGCTAAAAAGAAACAAGSAGAATTTAGATTAATAA	893
OY	1349	aacagcgttacgtacctgaagattcaacaacaatatagtggtgtgtgtggtc	1397
Db	894	GGCAATTGATCTTGATTAACCTCAACCAAAAATTAAGATTTGGCGAAGAT	942

RESULT 9
US-08-973-462-2
; Sequence 2, Application US/08973462B

```

: GENERAL INFORMATION:
: APPLICANT: DRUILHE, PIERRE
: APPLICANT: DAUBERSTES, PIERRE
: TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
: FILE REFERENCE: 0660-0125-0 PCT
: CURRENT APPLICATION NUMBER: US/08/973,462B
: CURRENT FILING DATE: 1998-02-06
: EARLIER APPLICATION NUMBER: PCT/FR96/00894
: EARLIER FILING DATE: 1996-06-12
: EARLIER APPLICATION NUMBER: FR 95/07007
: EARLIER FILING DATE: 1995-06-13
: NUMBER OF SEQ ID NOS: 29
: SOFTWARE: PatentIn Ver. 2.0

```

```

; LENGTH: 5361
; TYPE: DNA
; ORGANISM: P. falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(5361)
US-08-973-462-2

```

```

Query Match: 1.9%; Score 38; DB 4; Length 5361;
Best Local Similarity 44.9%; Prcd. No. 1.6;
Matches 266; Conservative 0; Mismatches 320; Indels 6; Gaps 3

```

OY	867	gtacltagtllatacaacaatttttaaaccgaatgtcgcaaaagccgatctggaaaaatt	926
Db	3699	tgtltaaatataatgcatactactaggcaaccttatttgaaaactcaagaagtlaatagaat	3758
OY	927	aettccaatacagaattagataatgtatgtatagtcgcaagaagaaaacaaatttggatta	986
Db	3759	agaagcagaatttataataagatatgtgaaaaatttaagaagatttagaaaagcatrtcaga	3818
OY	987	tatttacgaacctgatgycgaagaatltattatagataatttatgttgtcgttatttagatc	1046
Db	3819	agattctcaagaataataatagatgcanaaagatgatcacattagaaaaagtlattgaagaga	3878
OY	1047	tcaggtttaccaagcagcagtttgnaaaccttgcttcgagcaagccgcctgaagctgcgc	1106
Db	3879	acatgatataacgacgcagcgttggatgtagtattgtatgaatttaaaagtctcgaaagaca	3938
OY	1107	cataaagcgcagcaacagataacgaagyltaaccttaattaatgatitacagttagt---cta	1153
Db	3939	gaticgaataaagatctcgtattttaaaagtlcttgaaagaatataattataaagaaganaaga	3998
OY	1164	taacaaagccctctaagcaagatlatacaaatgaatttaaatgttaattgttcgcgtgcagc	1223
Db	3999	aalcaaaagaacttgaagytgaattttagaagatcataaagaatataaactatgtgaac	4058
OY	1224	agcaatttaacaatagagatcgtgatagtcgacaactgynaataatgttacaacatcogt	1283
Db	4059	agatattttgaagaagaa--aaaagaatagaaaaagtcatttgaaaaattcgaagaa	4116
OY	1284	gcggttattgscgttgaattcccacaagatgcagttaccaaaagtatagatgccttaat	1343
Db	4117	gaagctgaaagaataaaagaactcttgaagcagatatattataaagaagttcttcattcaa	4176
OY	1344	gttbaaacaggtttgacttgtaagttcaacaacacattaggttgytgytgaag	1395
Db	4177	gtcgaagaagaataaaattagaagaagtataacgaatttaaaagaagaggttag	4228
 RESULT 10 US-08-973-462-1			
; Sequence 1, Application US/08973462B			
; Patent No. 6191270			
; GENERAL INFORMATION:			
; APPLICANT: DRULHE, PIERRE			
; TITLE OF INVENTION: MALARIAL PRE-ERTHROCYTIC STAGE POLYPEPTIDE MOLECULES			
; FILE REFERENCE: 0660-0125-0 PCT			
; CURRENT APPLICATION NUMBER: US/08/973.462B			
; EARLIER FILING DATE: 1998-02-06			
; EARLIER APPLICATION NUMBER: PCT/FR96/00894			
; EARLIER FILING DATE: 1996-06-12			
; EARLIER APPLICATION NUMBER: FR 95/07007			
; NUMBER OF SEQ ID NOS: 29			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 1			
; LENGTH: 6152			
; TYPE: DNA			
; ORGANISM: P. falciparum			
; US-08-973-462-1			

Query Match	Similarity	1.9%	Score 38	DB 4	Length 6152
Best Local	Similarity 44.9%		Pred No. 1.7		
Matches 266	Conservative	0	Mismatches 320	Indels 6	Gaps 3
OY	808	ttaatgtgtcagtcacattcatagatgtacttataaaagg	-ggaagtga	ttgt	866
DB	3883	tgaaagtgttggaaggttttagaattgaaaataatc	tggaaggga	ttatgaagaatgc	3942
OY	867	gcatctagttctaacaaattctattacacagatgc	caaaagccg	atgtggaataatt	926
DB	3943	ttttgaaataatgcaattactctgcacaaattcttg	aagaatcc	gaggttlaattgaat	4002

[illegible]

RESULT 11
 US-07-991-867B-41/C
 Sequence 41, Application US/07991867B
 Patent No. 5476781
 GENERAL INFORMATION:
 APPLICANT: Moyer, Richard W.
 APPLICANT: Hall, Richard L.
 APPLICANT: Gruidl, Michael E.
 TITLE OF INVENTION: No. 5476781el Entomopoxvirus Expression System
 NUMBER OF SEQUENCES: 66
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: David R. Saliwanchik
 STREET: 2421 N.W. 41st Street, Suite A-1
 CITY: Gainesville
 STATE: FL
 COUNTRY: USA
 ZIP: 32606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentln Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/991,867B
 FILING DATE: 12-DEC-1992
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO 92/14818
 FILING DATE: 12-FEB-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/827,685
 FILING DATE: 30-JAN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/657,584
 FILING DATE: 19-FEB-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Saliwanchik, David R.
 REGISTRATION NUMBER: 31,794
 REFERENCE/DOCKET NUMBER: UF114.C3

APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanichik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF114.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8457 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE: Amsacta moorei entomopoxvirus
ORGANISM: Amsacta moorei entomopoxvirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (65..1459)
FEATURE:
NAME/KEY: CDS
LOCATION: 1474..2151
FEATURE:
NAME/KEY: CDS
LOCATION: complement (2239..2475)
FEATURE:
NAME/KEY: CDS
LOCATION: 2502..2987
FEATURE:
NAME/KEY: CDS
LOCATION: 3080..6091
FEATURE:
NAME/KEY: CDS
LOCATION: complement (6277..6768)
US-07-991-867B-1

Query Match 1.8%; Score 36.4; DB 1; Length 8457;
Best Local Similarity 53.2%; Pred. No. 5.2;
Matches 99; Conservative 0; Mismatches 86; Indels 1; Gaps 1;
Qy 841 tataaaagggagagatgctgtgattagttatacaaatattataacagatg 900
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7564 TATTAATATGGAAGTATTAATTTAAGCAATTAAGAAACATTAAGCTATATG 7505
Qy 901 -tcgcaaaagccagatcttggaataatccattaccagaattagataatgaaatt 959
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7504 TTCTTATTAAGATTATAGAACGAGATAGTTTAAAGAAACAAATTATATAGATG 7445
Qy 960 agcgaaagaaacaaagtgttgattatattacgaacctgacgaaagattataga 1019
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7444 TGCACAAAAATATATTATTCATATATATATATGTCAGATGAGCAATCTAAATTAT 7385
Qy 1020 taattt 1025
|||||
Db 7384 TAAATAT 7379

RESULT 14
US-08-544-332-1/c
Sequence 1, Application US/08544332
Patent No. 5935777
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5935777el Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 77
CORRESPONDENCE ADDRESS:
ADDRESS: Gerard H. Bencen
STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/544,332
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/991,867
FILING DATE: 07-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,755
FILING DATE: 19-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: NO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bencen, Gerard H.
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: UF114.C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8457 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Amsacta moorei entomopoxvirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (65..1459)
FEATURE:
NAME/KEY: CDS
LOCATION: 1474..2151
FEATURE:
NAME/KEY: CDS
LOCATION: complement (2239..2475)
FEATURE:
NAME/KEY: CDS
LOCATION: 2502..2987
FEATURE:
NAME/KEY: CDS
LOCATION: 3080..6091
FEATURE:
NAME/KEY: CDS
LOCATION: complement (6277..6768)
US-08-544-332-1

Query Match 1.8%; Score 36.4; DB 2; Length 8457;
Best Local Similarity 53.2%; Pred. No. 5.2;
Matches 99; Conservative 0; Mismatches 86; Indels 1; Gaps 1;
Qy 841 tataaaagggagagatgctgtgattagttatacaaatattataacagatg 900
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7564 TATTAATATGGAAGTATTAATTTAAGCAATTAAGAAACATTAAGCTATATG 7505

[illegible]

```

RESULT 15
US-08-323-1708-1
: Sequence 1, Application US/083231708
: Patent No. 5733772
:
GENERAL INFORMATION:
APPLICANT: Williamson, Kim C.
APPLICANT: Kaslow, David C.
TITLE OF INVENTION: Cloning and Expression of Plasmidium
TITLE OF INVENTION: falliparum Transmssion-Blocking Target Antigen, Pfs230
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
City: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/323,170B
FILING DATE: 13-OCT-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/010,409
FILING DATE: 29-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Quine, Jonathan A.
REGISTRATION NUMBER: P-41,261
REFERENCE/DOCKET NUMBER: 015280-11310005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 149..9556
US-08-323-1708-1

```

Query Match 1.8%; Score 36.2; DB 1; Length 9636;
Best Local Similarity 45.6%;
Matches 203; Conservative 0; Mismatches 238; Indels 4; Gaps 2

Dy	978	ttaggatataatttaacgaaccgatacgcgaagtattatattagataatttatgttcgctta	1037
Dy	1063	TGATGATTAGTGTTTTAAGAAGAAAGTGCGGGATGAACACGAAGAACATTATTCACAAGAGAAC	1122
QY	918	ggaaaattatcccttaccagaacttaagatatgcttgatttagcgaaagaacaagt	977
QY	1003	GGAAGAAGAGACGAGAGAGGAACAGGAGACGAAGAACGAAGAGAGACGACGAAGATA	1062

OY	1038	ttagaatctccaggttatatcaagcagcagttgaaaaacttgcttcttgcaagccgcgcg	1097
Db	1123	TCAGGAAGAAGTAGTGGTGCTGCAATCTTCACAAAGAAATTATTAATGATGGAGTGAAGATTCC	1182
OY	1098	aattgctgcacatgaagaagcgcgcaaccagataacgcaggttaacttaataagtgttaacgtt	1157
Db	1183	TGTAGGAAGCAGCGGGATGGAGATR---TGTAAAGCTTGACGAATTTMTAAGACCATAA	1239
OY	1158	agctataacaagaagccgcgtcaagcaagttatcaaatgtaattaatgaaatgtgtccgg	1217
Db	1240	TGGTGAACTTATGATGATAGTACATTAATAAATATGAAGTGTAGTGAAGAGTGAAGA	1299
OY	1218	tgcagcacgaatttaaccaatatagagg-alcggtlaatgycgaactggtgaaaatttgtacaat	1276
Db	1300	GGTAGGTGAAGAGGTGTAGTGCACAGAGGTAGTGCAGAGAGTGAAGAGTGTAGGTGAAGA	1359
OY	1277	cattcggttcggttatatgacgttgaattcccacaagaatgycagttaccaaaagtatatgatgc	1336
Db	1360	GGTAGGTGAAGAGAGTGTAGTGCACAGAGGTAGTGCAGAGAGTGAAGAGTGTAGGTGAAGG	1419
OY	1337	cttaaatgttgaagaacaggtttacta	1361
Db	1420	GGTAGGTGAAGAGTGTAGGTGAAGA	1444

Search completed: October 28, 2001, 15:36:00
Job time: 2239 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 14:58:41 ; Search time 1501.77 Seconds
(without alignments)
12412.704 Million cell updates/sec

Title: us-09-545-199C-3
Perfect score: 1972
Sequence: 1 agcggcgcattgcgtcagf.....gtctattcttcgttgataa 1972

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 segs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_est4: *
5: gb_est5: *
6: gb_est6: *
7: gb_est7: *
8: gb_est8: *
9: gb_est9: *
10: gb_est10: *
11: gb_est11: *
12: gb_est12: *
13: gb_est13: *
14: gb_est14: *
15: gb_est15: *
16: gb_est16: *
17: gb_est17: *
18: gb_est18: *
19: gb_est19: *
20: gb_est20: *
21: gb_est21: *
22: gb_est22: *
23: gb_est23: *
24: gb_est24: *
25: gb_est25: *
26: gb_est26: *
27: gb_est27: *
28: gb_est28: *
29: gb_est29: *
30: gb_est30: *
31: gb_est31: *
32: gb_est32: *
33: gb_est33: *
34: gb_est34: *
35: gb_est35: *
36: gb_est36: *
37: gb_est37: *
38: gb_est38: *
39: gb_est39: *
40: gb_est40: *
41: gb_est41: *
42: gb_est42: *
43: gb_est43: *

44: em_esthum10: *
45: em_esthum11: *
46: em_esthum12: *
47: em_esthum13: *
48: em_esthum14: *
49: em_esthum15: *
50: em_esthum16: *
51: em_esthum17: *
52: em_esthum18: *
53: em_esthum19: *
54: em_esthum20: *
55: em_esthum21: *
56: em_esthum22: *
57: em_esthum23: *
58: em_esthum24: *
59: em_esthum25: *
60: em_esthum26: *
61: em_esthum27: *
62: em_esthum28: *
63: em_esthum29: *
64: em_esthum30: *
65: em_esthum31: *
66: em_esthum32: *
67: em_esthum33: *
68: em_esthum34: *
69: em_esthum35: *
70: em_esthum36: *
71: em_esthum37: *
72: em_esthum38: *
73: em_esthum39: *
74: em_esthum40: *
75: em_esthum41: *
76: em_esthum42: *
77: em_esthum43: *
78: em_esthum44: *
79: em_esthum45: *
80: em_esthum46: *
81: em_esthum47: *
82: em_esthum48: *
83: em_esthum49: *
84: em_esthum50: *
85: em_esthum51: *
86: em_esthum52: *
87: em_esthum53: *
88: em_esthum54: *
89: em_esthum55: *
90: em_esthum56: *
91: em_esthum57: *
92: em_esthum58: *
93: em_esthum59: *
94: em_esthum60: *
95: em_esthum61: *
96: em_esthum62: *
97: em_esthum63: *
98: em_esthum64: *
99: em_esthum65: *
100: em_esthum66: *
101: em_esthum67: *
102: em_esthum68: *
103: em_esthum69: *
104: em_esthum70: *
105: em_esthum71: *
106: em_esthum72: *
107: em_esthum73: *
108: em_esthum74: *
109: em_esthum75: *
110: em_esthum76: *
111: em_esthum77: *
112: em_esthum78: *
113: em_esthum79: *
114: em_esthum80: *
115: em_esthum81: *
116: em_esthum82: *

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est110:*
191: gb_est111:*
192: gb_est112:*
193: gb_est113:*
194: gb_est114:*
195: gb_est115:*
196: gb_est116:*
197: gb_est117:*
198: gb_est118:*
199: gb_est119:*
200: gb_est120:*
201: gb_est121:*
202: gb_est122:*
203: gb_est123:*
204: gb_est124:*
205: gb_est125:*
206: gb_est126:*
207: gb_est127:*
208: gb_est128:*
209: gb_est129:*
210: gb_est130:*
211: gb_est131:*
212: gb_est132:*
213: gb_est133:*
214: gb_est134:*
215: gb_est135:*
216: gb_est136:*
217: gb_est137:*
218: gb_est138:*
219: gb_est139:*
220: gb_est140:*
221: gb_est141:*
222: gb_est142:*
223: gb_est143:*
224: gb_est144:*
225: gb_est145:*
226: gb_est146:*
227: gb_est147:*
228: gb_est148:*
229: gb_est149:*
230: gb_est150:*
231: gb_est151:*
232: gb_est152:*
233: gb_est153:*
234: gb_est154:*
235: gb_est155:*
236: gb_est156:*
237: gb_est157:*
238: gb_est158:*
239: gb_est159:*
240: gb_est160:*
241: gb_est161:*
242: gb_est162:*
243: gb_est163:*
244: gb_est164:*
245: gb_est165:*
246: gb_est166:*
247: gb_est167:*
248: gb_est168:*
249: gb_est169:*
250: gb_est170:*
251: gb_est171:*
252: gb_est172:*
253: gb_est173:*
254: gb_est174:*
255: gb_est175:*
256: gb_est176:*
257: gb_est177:*
258: gb_est178:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	327.2	16.6	616	236	AQ989871	RfC00534
2	216.8	11.0	334	108	AU175106	AU175106
3	200.8	10.2	626	136	BE496766	BE496766
4	199	10.1	606	104	AT975327	AT975327
5	198.2	10.1	699	30	AV404766	AV404766
6	196.8	10.0	820	106	AU000654	AU000654
7	187.6	9.5	746	106	AU000653	AU000653
8	185.8	9.4	982	106	AL533571	AL533571
9	184.6	9.4	918	105	AL514744	AL514744
10	184.6	9.4	978	105	AL518359	AL518359
11	183.8	9.3	914	106	AL541312	AL541312
12	183.4	9.3	931	106	AL536748	AL536748
13	183.4	9.3	985	105	AL516904	AL516904
14	183.4	9.3	985	105	AL518879	AL518879
15	183.4	9.3	1047	105	AL514536	AL514536
16	183	9.3	938	106	AL517232	AL517232
17	182.8	9.3	908	105	AL517232	AL517232
18	182.6	9.3	873	106	AL539713	AL539713
19	182.6	9.3	931	105	AL515529	AL515529
20	182.6	9.3	932	106	AL540455	AL540455
21	182.6	9.3	932	106	AL560427	AL560427
22	182.6	9.3	953	106	AL539740	AL539740
23	182.6	9.3	964	106	AL557006	AL557006
24	182.6	9.3	983	106	AL557761	AL557761
25	182.6	9.3	997	106	AL538946	AL538946
26	182.6	9.3	997	106	AL559949	AL559949
27	182.6	9.3	1033	105	AL515945	AL515945
28	182.6	9.3	1038	106	AL516836	AL516836
29	182.4	9.2	879	106	AL560200	AL560200
30	182.4	9.2	975	106	AL539067	AL539067
31	182.2	9.2	947	141	BE908742	BE908742
32	181.8	9.2	673	164	BE205335	BE205335
33	181.8	9.2	936	105	AL517801	AL517801
34	181	9.2	1080	105	AL515288	AL515288
35	179.8	9.1	1038	106	AL557188	AL557188
36	179.2	9.1	893	105	AL516060	AL516060
37	179.2	9.1	1101	137	BE543866	BE543866
38	178.8	9.1	800	14	AA948879	AA948879
39	178.6	9.1	439	22	AT574511	AT574511
40	178.6	9.1	919	105	AL515004	AL515004
41	178.4	9.0	868	105	AL519001	AL519001
42	177.2	9.0	798	146	BF259735	BF259735
43	177	9.0	982	105	AL516816	AL516816
44	176.8	9.0	883	106	AL541791	AL541791
45	176.4	8.9	890	106	AL560126	AL560126

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
1	AQ989871/c	616 bp DNA	AQ989871	GI:9648465	GSS.	Photorhabdus luminescens	Photorhabdus luminescens	1 (bases 1 to 616)	french-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T., Daborn, P., Bowen, D., and Blattner, F.R.	A genomic sample sequence of the entomopathogenic bacterium Photorhabdus luminescens w14: potential implications for virulence
		RfC00534 Photorhabdus luminescens strain w14 M13 library								
		Photorhabdus luminescens genomic clone PLG00534, DNA sequence.								
		GSS								
		14-AUG-2000								
		Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae; Photorhabdus.								

JOURNAL
MEDLINE
20378633
COMMENT

Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)
Contact: french-Constant RH
Department of Biology and Biochemistry
University of Bath
South Building, Bath BA2 7AY, UK
Tel: (44) 1225 826621
Fax: (44) 1225 826779
Email: bsr@bath.ac.uk

This is one of 2,122 random reads from the M13 library. For annotation of identified clones (BLASTX, BLASTN and mapping to E. coli K12 genome) please see french-Constant et al. 2000, Nucleic Acids Res.
Seq primer: M13 Forward
Class: Shotgun.
Location/Qualifiers
1. 616
/organism="Photorhabdus luminescens"
/strain="w14"
/db_xref="taxon:29488"
/clone="PLG00534"
/clone_lib="Photorhabdus luminescens strain w14 M13 library"
/dev_stage="primary phase variant"
/note="Genomic DNA from strain M14 was size selected (1-2 kb) and then cloned into M13 Janus."

BASE COUNT 173 a 164 c 116 g 151 t 12 others
ORIGIN

Query Match 16.6% Score 327.2; DB 236; Length 616;
Best Local Similarity 76.7% Pred. No. 9.4e-69;
Matches 442; Conservative 0; Mismatches 130; Indels 4; Gaps 4;

Qy	1399	gctgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	1457
Db	616	gctgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	557
Qy	1458	aatacgaatcgtatcgtatgagatgagacgaacacatggtgctatgatacgtatg	1517
Db	556	gctgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	497
Qy	1518	ggtgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	1577
Db	496	ggtgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	438
Qy	1578	ggtgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	1637
Db	437	ggtgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	378
Qy	1638	ggtgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	1697
Db	377	ggtgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	319
Qy	1698	ggtgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	1756
Db	318	ggtgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	259
Qy	1757	ggtgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	1816
Db	258	ggtgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	199
Qy	1817	ggtgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	1876
Db	198	ggtgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	139
Qy	1877	ggtgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	1936
Db	138	ggtgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	79
Qy	1937	ggtgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	1972
Db	78	ggtgatacgaatgagatcatcgtatgataaagcgcggttaacgcgtataca-atacg	43

```

RESULT 2
LOCUS   AUI75106      334 bp      mRNA      EST      21-MAR-2001
DEFINITION AUI75106 Rice panicle at ripening stage Oryza sativa cDNA clone
            E10189, mRNA sequence.
ACCESSION AUI75106
VERSION   AUI75106.1  GI:13421016
KEYWORDS  EST.
SOURCE    Oryza sativa.
ORGANISM  Oryza sativa
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
           Ehrhartioideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 334)
AUTHORS   Sasaki, T. and Yamamoto, K.
TITLE     Rice cDNA from panicle at ripening stage (2001)
JOURNAL   Unpublished (2001)
COMMENT   Contact: Takuji Sasaki
           National Institute of Agrobiological Resources
           Rice Genome Research Program
           2-1-2 Kannondai, Tsukuba
           Ibaraki,
           Japan 305
           Tel: 0298-38-7441
           Fax: 0298-38-7468
           Email: tsasaki@abr.affrc.go.jp
PROJECT   " RGP.
E10189_62.

FEATURES
Source    1..334
           /organism="Oryza sativa"
           /strain="Nipponbare"
           /db_xref="taxon:4530"
           /clone="E10189"
           /clone_1lb="Rice panicle at ripening stage"
           /dev_stage="ripening stage"
           /note="Organ: panicle; Rice cDNA from panicle at ripening
           stage"

BASE COUNT      83 a      76 c      97 g      78 t

ORIGIN
Query Match      11.0%; Score 216.8; DB 108; Length 334;
Best Local Similarity 78.3%; Pred. No. 4.5e-42;
Matches 260; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

OY 1535 gcaaggtgaatcggtgcagaagaagtgtctattacccgtgcgcaacaagtatga 1594
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   3 GAAAGCGAGATCGGTGAAAGAGCGCTTGGCGCATTCACCGCAGCACCTTCCTACGA 62

OY 1595 agaacaattcaacgactagacttttagaaacgggaattaaagtacgactagttg 1654
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   63 AGAGCTGTCAACACTCTCAGAACTGCTGGAACCGGTATCAAGTTATCGACTGATGTG 122

OY 1655 tccggttcggaagaagggttaagtgttatttcggtgtgcggtgtcgtttaaaccgt 1714
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  123 TCCGTTCCGTTAAGGCGGTTAAAGTGTGCTGTGCGTGTGCGGTGTAAGTAAACCGT 182

OY 1715 caatatgtgaattaatcgcgaatacgcgaattgaagcaactcgaactctgtcttgc 1774
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  183 AAACATGATGAGGAGTCATTCATCATCGATCGCATCGACATCCGGTTACTCTGTGTTGTC 242

OY 1775 gggggatggtgagcgtgaagcgtgaagcgtgaagcgttatacgtatgtaagaagacttaa 1834
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  243 GGGCGTACGATGACCTACTCGTGAAGGTAACGACTTCTACACGAATGACCGACTCCAA 302

OY 1835 cgtatagataaagtgctctgttattatgctc 1866
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   303 CATTATCGACAAAGTATCTCCTGCTATGACC 334

```

```

RESULT 3
LOCUS   BE496766      626 bp      mRNA      EST      03-AUG-2000
DEFINITION BE496766 Hc_d11_08F03_SKPL Haemonchus contortus d11 mixed adult from David
            Knox Haemonchus contortus cDNA clone Hc_d11_08F03 5' similar to
            sPLP46561ATPB.CABEL PROBABLE ATP SYNTHASE BETA CHAIN,
            MITOCHONDRIAL PRECURSOR, mRNA sequence.
ACCESSION BE496766
VERSION   BE496766.1  GI:9690101
KEYWORDS  EST.
SOURCE    Haemonchus contortus.
ORGANISM  Haemonchus contortus
           Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdita; Strongyloida;
           Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.
REFERENCE 1 (bases 1 to 626)
AUTHORS   Blaxter, M.L., Parkinson, J., Whilton, C., Daub, J., Gilliano, D., Hall
           N., Quayle, M., and Barrell, B.
TITLE     Edinburgh University/Sanger Centre Nematode EST Project
JOURNAL   Unpublished (2000)
COMMENT   Contact: Blaxter ML
           Institute of Cell, Animal and Population Biology
           University of Edinburgh
           Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
           3JF, UK.
           Tel: +44 131 650 6760
           Fax: +44 131 670 5450
           Email: mark.blaxter@ed.ac.uk
           The library was prepared by David Knox of the Moredun Institute,
           Edinburgh. Sequencing was performed by Claire Whilton, ICAPB,
           Edinburgh
           PCR Primers
           FORWARD: T3
           BACKWARD: T7PL
           Plate: 08 row: F column: 03
           Seq primer: SKPL
           High quality sequence stop: 547.
           Location/Qualifiers
           1..626
           /organism="Haemonchus contortus"
           /strain="Moredun"
           /db_xref="taxon:6289"
           /clone="Hc_d11_08F03"
           /clone_1lb="Haemonchus contortus d11 mixed adult from
           David Knox"
           /sex="Mixed"
           /dev_stage="Adult, day 11"
           /note="Vector: Lambda Zap II; Site_1: EcoRI; Site_2: XhoI;
           Constructed by David Knox, Moredun Institute, Edinburgh.
           Primary library titre was 4x10^6 pfu."

BASE COUNT      157 a      149 c      179 g      141 t

ORIGIN
Query Match      10.2%; Score 200.8; DB 136; Length 626;
Best Local Similarity 61.4%; Pred. No. 3.6e-38;
Matches 350; Conservative 0; Mismatches 202; Indels 18; Gaps 1;

OY 1351 caggtttagcttgaagttcaacaacaattagtggtgtgtgtgttcgtcgtatcgaa 1410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   38 CACGCTTATCTCCTCGAATCTCTCAGCATCTAGGACAAATGTGGTCCGTTGCAATGCTA 97

OY 1411 tgggatctctgtagtgattaaacgcggtttaagcgtaacaatacgsataccaattt 1470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   98 TGGACGGAACTGAGGGACTTGTTCGGGCTCAAGAAAGTTGTGACACCTGATCAATCA 157

OY 1471 cgttcacgttgggaacgaacaacattgggttcgtatcatgaacgtatttgggtgaaccaatcg 1530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   158 AGATTCCAGTAGGCGCAGAGACTCTCGGCGCAATATGAAATGTCATCGCGCAACCAATCG 217

OY 1531 atgagcaagtgtaaatcgtgtgcagaagaagaatttggtctattcaacggtgcgcaacaatt 1590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   218 ACGAGCGAGGTCCTATCCCCCGCAGCACCGCGCTCTATTCACCTGAGGACCAAGAGT 277

```

[illegible]

RESULT	4
LOCUS	A1975327
DEFINITION	A1975327 606 bp mRNA EST 27-AUG-1999
ACCESSION	S81269922 Schistosoma mansoni female, Phil Love/Dee Merrick
VERSION	Schistosoma mansoni cDNA clone SMrAR4 5' end, mRNA sequence.
KEYWORDS	A1975327 A1975327.1 GI:5788495
SOURCE	EST.
	Schistosoma mansoni.

Eukaryota; Metazoa; plathyhelminthes; Rhaditophora; Neodermata; Trematoda; Digenea; Strigeidida; Schistosomatoida; Schistosomatidae; Schistosoma.

REFERENCE	1 (bases 1 to 606)
AUTHORS	Merrick,J.M., Osman,A., Loverde,P.T., Chandra,I., Glodex,A., Fraser, C.M. and Lee,N.H.
TITLE	Generation of a Schistosoma sp. EST (SMEST) Catalog & Schistosome Gene Index
JOURNAL	Unpublished (1998)
COMMENT	Contact: Norman H. Lee

COMMENT

Contact: Norman H. Lee
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel.: 301 838 3529
Fax: 301 838 0208
Email: nhlee@tigr.org
Seq primer: M13 Reverse.

FEATURES	Location/Qualifiers
source	1. .606

```

/organism="Schistosoma mansoni"
/db.xref="taxon:6183"
/clone="SMFAR46"
/clone_id="Schistosoma mansoni female, Phil Lloverde/Joe
Merrick"
/sex="female"
/note="Vector: plasmidscript SK; Site_1: EcoRI; Site_2: XhoI
; directionally cloned cDNAs"
164 a 121 c 159 g 162 t

```

Query Match	10.1%	Score 199	DB 104	Length 606
Best Local Similarity	60.4%	Pred. No 9	7e-38	
Matches 356	Conservative 0	Mismatches 215	Indels 18	Gaps 1
Oy	1354	gcttagtaactggaagttcaacaacaaattagtggtggtgtagttctgcgtatcgcgaatg	1413	
Db	18	gctcttatattggaagttccacacaacattctgttggtggaacaacactgtgtggcaccacattggtatgac	77	

QY 1414 gacatcctgtagtgattaaacgcggttttaagcgtcaacaataagaaataaaccattttcgt 1473
 Db 78 ACGGTACCGAAGGTCCTAGTTCCTGGCGGTAGCAGAGTGTGTTGATACGAGCGGTCTCAATTCGTA 137
 QY 1474 ttccagctgggaacgaaacaatttgctgcctatcatcgaacgctattggctgaaaccaatgatg 1533
 Db 138 TTCCGTGTTGGTCTTGAGACCTTGTGGCCGTATCATGAAACGTTATTCGGGAACCAATCGATG 137
 QY 1534 agcaagctggaatctggtgcagaagaagaatttgctcattaccgtgcgcaccacaagtatg 1593
 Db 198 AGAGAGACCCATCAAAACGAAATGATGTGGGTATTCAACACAGAGATGCTCCAGACTTCA 257
 QY 1554 aagaacaactcaacagctactggaacttttagaaccgggaatlaaagtatcgaactggtt 1653
 Db 258 TTGGAATAGTACTGAGACAGAAATTTTACGACGAGCATCAAGGTTGTGAGACTGTATG 317
 QY 1654 gtccgcttgcgaagaagggtgaagtagttattcggctggtgcggtgctgtaaaacg 1713
 Db 318 CTCCTTATGTGAAGGAGGAAATATTCGCTTTTCGGGTGGTGGGGGTTGGCAAAATCG 377
 QY 1714 tcaatctatgtaaatatccgtacacgcgaatttgacacgcactcaggtactcgtcttg 1773
 Db 378 TACTTATTATGAGACTATCAATAATATTGCGACAGGCTACAGGAGGTACTCTGTGTTTG 437
 QY 1774 cgggggtgaagtgcgtacgcgctgtaagtgtaacgactctatcatalgatagtaaagactcta 1833
 Db 438 CTGGGTGTTGGTGAACGTAACCTCGAAGAGCAATGATTTATTCAGAAATGATTAACACTG 497
 QY 1834 acgtattaga-----taagtgtctcttgtttatggtccaatgaag 1875
 Db 498 GTGCGATTGACCTTAAAGAAACAATTCCTAAGGCTGCTCGGTGTACGGGCAATGATG 557
 QY 1876 agccacagctgaaccgctttagcgttgccattaaacaagctaaactatgcg 1924
 Db 558 AACCAACAGAGGACAGCTGCTGTGTTGCTCTTAACGGGTCTTAACCAATTGC 606

RESULT	5
LOCUS	AV404766
DEFINITION	AV404766 699 bp mRNA EST 06-FEB-2000 AV404766 <i>Bombyx mori</i> prothoracic gland p50 5'h-instar day 4 larva <i>Bombyx mori</i> cDNA clone ppgv0416 T3, mRNA sequence.

ACCESSION	AV404/56	
VERSION	AV404/66.1	GI:6908854
KEYWORDS	EST.	
SOURCE	domestic silkworm.	
ORGANISM	<i>Bombyx mori</i>	

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 699)
Mita, R., Moriyama, M., Shimada, T., Okano, K. and Maeda, S.
Bombyx mori cDNA
Unpublished (2000)
Contact: Mita K

Euarystota, Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota, Neoptera, Endopterygota; Lepidoptera; glossata; Ditrysia
; Bombycoidea; Bombycidae; Bombyx.

Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: km1ta@nirs.go.jp
method: uni-directional, sequence direction: sequenced from 73 primer
(5' -> 3')
Project- 'Silkworm Genome Program in MAF, and Research for the
Future Program in JSPS'. see: Silkworm . see: Silkworm .
<http://www.ab.a.u-tokyo.ac.jp/silkworm/>, for whole ESTdb.
Location/Qualifiers
1. .699

```

/organism="Bombyx mori"
/strain="p50"
/db_xref="taxon:7091"
/clone="p19v0416"
/clone_1b="Bombyx mori prothoracic gland p50 5th-instar
day-4, larva"
/sex="Female/male mixed"

```

BASE COUNT	166 a	161 c	178 g	193 t	1 others
ORIGIN	/tissue_type="prothoracic gland" /dev_stage="5th-instar-day-4 larva"				
Query Match	10.1%; Score 198.2; DB 30; Length 699;				
Best Local Similarity	60.1%; Pred. No. 1.5e-37;				
Matches 357; Conservative	0; Mismatches 219; Indels 18; Gaps 1;				
Qy 1395	gttcgcgtatcgcgaatggatcatctcgtatggatlaaagcggtttaacgctaacaaat 1454				
Db 8	gtytcgacacattggccaaaggacggcttactcgttgggacacccgctactagac 67				
Qy 1455	acgaataaccacattctcttcagctggagacgaacacattggctgatalcagaacta 1514				
Db 68	tctgcctcaccatttcgtatcccgctggagctgaamaacctcggtccatcatcaatga 127				
Qy 1515	ttgggtgacccaatcgaatgagcaagctgaactcggctgcaagaagaattgycatcac 1574				
Db 128	atcgccgaacacgattgacgagcatttccatccaccgacacagactcgtgatttcac 187				
Qy 1575	cgtgcgcaccaagctatgaagaacaactcaacagctctgtaactttagaacoggaatt 1634				
Db 188	gctgaacgtccagagtttgcgacatgctgtgcacgacgagatttcggaactggtata 247				
Qy 1635	aaagtatcagcttaagtttgcgcttggcaagaagggttaagtagttatctcgtcgt 1694				
Db 248	aaagtcgtcgtatcgtcgtcgtcgttattgccaagagagaaagattgggttcttggcgga 307				
Qy 1695	gcgggttcggtlaaaccgctcaatatgattgaattaaatccgtaacatcgcaattgagac 1754				
Db 308	gctggtgtgggcaaaacgattgatttatggaactgatacaacattgtggcaaacccat 367				
Qy 1755	tcaggttaactctgtcttggcggggtgagtgagctgacgctgtagaagtgaagactctat 1814				
Db 368	ggtggttactctgtgtgtgtgctgagtagagagcgtactcctgaaggaagattttatnc 427				
Qy 1815	catgagatgaaagactct-----aacgtattagaatagtgctctt 1856				
Db 428	caccgattgattgattcgtgctgattttcttaaaagcacaacatccaaagtagctcta 487				
Qy 1857	gttatgtgcaaatgaacagacccacacagtaacggttaacgttgatgtgacatcaacagctta 1916				
Db 488	gtatttggtcaaatgaacagac 547				
Qy 1917	actatgctgcggaataatccgtgtagaagaagctgtagtgcctatcttcgttgat 1970				
Db 548	actgttgcctggaattttccgtagatcaaaagacacagatgactcctttcat 601				
RESULT 6					
AU000654	LOCUS 820 bp mRNA EST 15-JAN-1999				
DEFINITION	AU000654 Bombyx mori p50(Daizo) Bombyx mori cDNA clone e40837, mRNA sequence.				
ACCESSION	AU000654				
VERSION	AU000654.1 GI:4156898				
KEYWORDS	EST.				
ORGANISM	domestic silkworm.				
SOURCE	Bombyx mori				
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea; Bombycidae; Bombyx.				
AUTHORS	Mita,K., Morimyo,M., Shimada,T., Okano,K. and Maeda,S.				
TITLE	Establishment of cDNA database of Bombyx mori				
JOURNAL	Unpublished (1999)				
COMMENT	Genome Research Group National Institute of Radiological Sciences Anagawa 4-9-1, Inage, Chiba 263-8555, Japan Email: kmite@nirs.go.jp				

FEATURES		PROJECT = 'CRESP project by JST'.	
SOURCE		Location/Qualifiers	
		1..820	
		/organism="Bombyx mori"	
		/strain="p50(Daizo)"	
		/db_xref="taxon:7091"	
		/clone="e40837"	
		/clone_1lb="Bombyx mori p50(Daizo)"	
BASE COUNT	210 a 178 c 219 g 213 t		
ORIGIN			
Query Match	10.0%; Score 196.8; DB 106; Length 820;		
Best Local Similarity	57.5%; Pred. No. 3.4e-37;		
Matches 407; Conservative	0; Mismatches 277; Indels 24; Gaps 2;		
Oy	1220 cagcagcaattacaacatagaggatcggtaatgccaactgynaanaattgatacaatcat	1279	
Db	109 CAGAGGGACATATGCTGCTAATATCGCAGGAAAGGCCAAGGTAAAGGTGTTGCCGTTAT	168	
Oy	1280 cggltgcggttatgacgttgaattcccaaga-----tgcagtlaccaaaagtatatga	1333	
Db	169 CGGTCTGTGGTAGATGTGCAATTCGAAGATTAACCTGCCCTCGATTCTTAATGTCCCTTGA	228	
Oy	1334 tgccttaaatgttgaaagaagtttagcttagaagttaacaaacaatatggtgtgtgt	1393	
Db	229 GGTGCAAAATCGATCTCCCGCCCTCGTACTTGAGGTAGCCCAACATTGGGTGGAACAC	288	
Oy	1394 agttgcgcgtatcgcaatcggagatcatcgtatgataltaaacgcggtttaagcgtaacaa	1453	
Db	288 AGTTGCGACATATGTCATGCAATGCAAGCGTACTGAAGCGCTACTCGTGGCAACCGCTACTGGA	348	
Oy	1454 taacgaataaccaattctgttccagttgagtggaacgaanaaatltggtcatcgtgaacgt	1513	
Db	349 CTCTGCTCACCAAATTCGTATCCCGTGGAGCTGAAACCTCGCTGCATCATCAATGT	408	
Oy	1514 atgtgttgaaaccaatcgtatgagcaagtggaatcgtgtcagaagaagaattggtctatca	1573	
Db	409 AATCGGCGCAACGATTTGACGAGCGTGTGCCATCCCAACGACGACACTGCTGCTATTCA	468	
Oy	1574 ccgtgcgcacccaagttatgaaacaactctaacagtaactgaactttagaacgcggaat	1633	
Db	469 TCGTGAAGCTCCAGAGTTTGTGCAATGTCTGTGCAAGCAGAGAGATTCTCGTAACGTGAT	528	
Oy	1634 taaagtltatcgaacttagttgtcgtttgcgaaggggtaagtagttatcctcgtgtg	1693	
Db	529 AAAAGTCGTCGATCTGCTGCTCTTATGCCAAAGAGAAAGATTGGGTTGTCGGG	588	
Oy	1694 tgcggtgtgcgttaaacgcgtcaatataigtatgaattatcgttaacatcgcgaattgagca	1753	
Db	589 AACTGCTGTGGGCCAAAACGTAATTGATATAGAACCTGATCAACAATGTTGCCAAAGCCCA	648	
Oy	1754 ctcaagttactctgtccttgcgggggtagttagcgtacgcgtgaaggtaacgactcta	1813	
Db	649 TGTGTGTTACTGTGTGTTTGTGCTGAGATGAGGAAGCGTACTGTGAAGAAATGATTATTA	708	
Oy	1814 tcatgagatgaagaactc-----aacgtatataagaagtgtcct	1855	
Db	709 CCAACGATGATGATGATCTGCTGATATTTCTTAAGACCAAAACATCCAAAGTGTCT	768	
Oy	1856 tgtttatgttcaaatgaagagcagccacgaagtatacgtttacgtgtgc	1903	
Db	769 AGTATGTGTCAAATGACGAACCTCTGTGTCCTGCTGCTGTGTC	816	
RESULT	7		
AUD000653	746 bp	mRNA	EST
LOCUS	AUD000653	Bombyx mori p50(Daizo)	Bombyx mori cDNA clone e40836, mRNA
DEFINITION	sequence.		
ACCESSION	AUD000653		
VERSION	AUD000653.1	GI:4156897	
KEYWORDS	EST.		

SOURCE domestic silkworm.
ORGANISM Bombyx mori
REFERENCE Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia ; Bombycoidea; Bombycidae; Bombyx.
AUTHORS 1 (bases 1 to 746)
TITLE Mita,K., Morimyo,M., Shimada,T., Okano,K. and Maeda,S.
JOURNAL Establishment of cDNA database of Bombyx mori
COMMENT Unpublished (1999)
CONTACT: Mita K
Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmits@nirs.go.jp
PROJECT = 'CREST project by JST',
Location/Qualifiers
1. 746
/organism="Bombyx mori"
/strain="p50(Daizo)"
/db_xref="taxon:7091"
/clone="e40836"
/clone_11b="Bombyx mori p50(Daizo)"
BASE COUNT 191 a 159 c 202 g 194 t
ORIGIN

Query Match 9.5%; Score 187.6; DB 106; Length 746;
Best Local Similarity 57.5%; Pred. No. 5.7e-35;
Matches 359; Conservative 0; Mismatches 259; Indels 6; Gaps 1;

1220 cagcgcaatttaacaataagagatcgtaatgcaactgcaaaatgtaacaatc 1279
109 CAAGAGGAGACTGCTGCTAAATCTGCAGAGAAAGCCCAAGTTAGTGTTCCTTAT 168
1280 cgggtcggtatgacgttgaaatcccaaga-----tgcaatccaaagtatga 1333
169 CGGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 228
1334 tgccttaattgtgaacaggtttgacttgaaattcaacaacaaattggtgtgtgt 1393
229 GGTGCAAAATCATCTCCCGCCCTGCTGATGAGTACGCCCAACATTTGGTGAGAACAC 288
1394 agtgcgtatcgatcgatgagatcgtatgattaaagcggttaagcgtaacaa 1453
289 AGTTGGACACATTCGATGAGAGGATGATGATGATGATGATGATGATGATGATGAT 348
1454 taagaataaccgaatttcgttcagtggaacgaacaaattggtgtgtatcatga 1513
349 CTCTGGTACCAATTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 408
1514 attggtgtaacgaatcgatgagaaatcgatcgatcgatcgatcgatcgatcgat 1573
409 AATCGGCAACGATGAGAGGATGATGATGATGATGATGATGATGATGATGATGAT 468
1574 cgttcgacacgaatttgaagaacaaatcgaactgaactgaacttgaagaagggaat 1633
469 TGCTAAGCTCCAGATTTGTCGACATGCTGTGACAGCAGAGATTTCTGTAATGCTAT 528
1634 taaagtacgaactagttgctgctgttcgaagaagggttaagaagttaattcgatg 1693
529 AAAATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 588
1694 tgcggtgttcgttaaacgcgtcaataatgataatgaatcgtaacatcgcaatgagca 1753
589 AGCTGCTGTGGGCAAACTGATTTGATGATGATGATGATGATGATGATGATGATGAT 648
1754 ctcaagtaactctgttcgttcggtggtggtggtggtggtggtggtggtggtggtg 1813
649 TGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 708
1814 tcatgagatgaagaactcaact 1837
709 CCACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 732

RESULT 8
AL533571 982 bp mRNA EST 13-FEB-2001
LOCUS AL533571 LTI_FL013_Fbrn1 Homo sapiens cDNA clone CS0DF002Y109 5
DEFINITION prime, mRNA sequence.
ACCESSION AL533571
VERSION AL533571.1 GI:12797064
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 982)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 982
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DF002Y109"
/clone_11b="LTI_FL013_Fbrn1"
/dev_stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"
/lab_host="DH10B"
/note="organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville , Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifestech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 224 a 223 c 281 g 249 t 5 others
ORIGIN

Query Match 9.4%; Score 185.8; DB 106; Length 982;
Best Local Similarity 55.0%; Pred. No. 1.6e-34;
Matches 389; Conservative 4; Mismatches 296; Indels 18; Gaps 1;

1284 gcggttattgacgttgaaatcccaagaatgcaatcaaaagtatatgatgcctta 1343
207 GTGGACGTCACATTTGATGAGGAGGACTACCAATTCATAATGCCCTGGAAGTCAG 266
1344 gtggaacaggttgatgacttgaaatcaacaacaaattggtgtgtgtgtgtgtgt 1403
267 AGGAGACCAAGACTGTTTGGAGGTGGCCAGCATTTGGGTGAGACCAATAAGGAT 326
1404 atcgcaatggatcatctgattgaaatgaacgcgttgaaatgaacaaatgaatga 1463
327 AATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 386
1464 ccaatttcgttcagtggaacgaacaaatggtgtgtgtgtgtgtgtgtgtgtgtgt 1523
387 CCAATCAAAATTCCTGTTGCTCTGATGATGATGATGATGATGATGATGATGATGAT 446
1524 ccaatcgatgaacgaaggtgaatcggtgcgagaagaattgtgtcatatccacgtgc 1583
447 CCTATTGATGAAGAAGRGTCATCAAAACCAATTTGCTCCATTTATGATGATGAT 506
1584 ccaatltatgaagaacaaatcaacagtaactttagaagcgggaattaaatltac 1643
507 CCAGAGTTCATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 566

RESULT	9
AL514744	
LOCUS	AL514744 918 bp mRNA EST 13-FEB-2001
DEFINITION	AL514744 LRT_NFL006_Pf2 Homo sapiens cDNA clone ClDB0B13ZH09 5
ACCESSION	prime_mRNA sequence.
VERSION	AL514744
KEYWORDS	AL514744.1 GI:12778237
SOURCE	EST.
ORGANISM	human.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 918)
AUTHORS	I.I.,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 101 91006 Evry cedex - France Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Query Match	9.48;	Score 184.6;	DB 105;	Length 918;
Best Local Similarity	55.08;	Pred. No. 3.1e-34;		
Matches 389;	Conservative 3;	Mismatches 297;	Indels 18;	Gaps 1

OY 1284 gcggttatgcagcttgaatcccaacagatgagtaccaaagtatatgatgccttaaat 1343

DB 206 GTGGACGCTCAGTTTGATGAGGAGACTACMACCATTTTAAATGCCCTGGAGTCACAGGC 265

RESULT	10
AL518359	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNML	
COMMENT	

AL518359 978 bp mRNA EST 13-FEB-2001
 AL518359 LTI_NF011_NBC1 Homo sapiens cDNA clone CS0DA009YB03 5
 prime, mRNA sequence.
 AL518359
 AL518359.1 GI:12781852
 EST.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 978)
 Li, W.-B., Gruber, C., Jesssee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: secrete@genoscope.cns.fr, web : www.genoscope.cns.fr.

```
FEATURES
  source
    location/Qualifiers
      1..978
        /organization="Homo Sapiens"
        /db_xref="taxon:9606"
        /clone="CS0DA009YB03"
        /clone.lib="LTI-NFLO11_NBC1"
        /sex="male"
```



```

/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(37) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. library is not normalized, but is the control for
the normalized libraries. library was constructed by Life
Technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville,
Maryland 20850, USA Fax : (41) 301 610 8371 Email :
fliang@lifestech.com URL :
http://fulllength.invitrogen.com"

```

Query Match	9.4%	Score 184.6	DB 105	Length 978
Best Local Similarity	55.0%	Pred. No. 3.1e-34		
Matches 389	Conservative 3	Mismatches 297	Indels 18	Gaps 1

[illegible]

RESULT	11
AL541312	
LOCUS	914 bp mRNA EST 16-FEB-2001
AL541312	

DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES
AL541312 tLTL_F1002_PL1 Homo sapiens cDNA clone GS0DE006117 5 prime	AL541312	AL541312	EST.	human.	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	L1,W.B., Gruber,C., Jesssee,J. and Polayes,D.	Full-length cDNA libraries and normalization	Unpublished (2001)	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: segret@genoscope.cns.fr, Web : www.genoscope.cns.fr.	Location/Qualifiers 1..914

	BASE COUNT	ORIGIN
	214 a	205 c 268 g 226 t 1 others

Query Match	9.3%	Score 183.8;	DB 106;	Length 914;
Best Local Similarity	55.5%;	Pred. No. 4.9e-34;		
Matches 412; Conservative	0;	Mismatches 307;	Indels 24;	Gaps 2;

OY	1254	gcaacttgaaaaatgfracaacatccgfrgsggttattgcagtgtgaatcccaaat	1313
Db	170	gccaccgggcacatccgsgcgctattggcgacgtgtgacgtccacttgatgagsga	229
OY	1314	gcagtaccaaaagtatat-----gatgccttaaatgttgaaacaggtttagtactgaa	1367
Db	230	ctaccaccacttctcaatggccctgaaagtgacaaagcagagaccagactggTTTTGGAG	289
OY	1368	gttcaacaacaatagtggtgtgtgtagtgcgtatcgcatacgaatggatcatctga	1427
Db	290	gtggccccagcatttgggtgtgagacacagtaagagactattgcatggaatggtacaaagc	349
OY	1428	ttaaacgcggtttaaagcgtfaacaacaaatagcaataaccacattcgtccagttggaacg	1487
Db	350	tttggttagagccacgaagaatgactgagattcttggtgcacacataaaatttcttggctct	409
OY	1488	aaaacattggtgtcgtatcatgtaaacgfatattggtgaaaccaatcgatgagcaagtgaac	1547
Db	410	gagactttgggcacaaatgcatagaattgctatggagacactattgtaaaagaggtcccatc	469
OY	1548	ggtgcagaaggaattgtctattcaaccgtgtgcaccacgaagtatagaagaacaatcaac	1607
Db	470	aaaacaaacatttgcctccattctacgtgacgctccagactgacgtatcagtaaaatgagttt	529
OY	1608	agtaactgaactttagaacaagggaatlaaagtatacgaacttattgttcggtttgcgaa	1667
Db	530	gagcagagaaatttcgtgacatcgatgatacgaagttgtgcattctgtagctccatgccaag	589
OY	1668	gggggttaagtagttattcgtgtgtgtgcggtgtgtcgtgtaaaacgcgtacaatatgatgaa	1727
Db	590	ggtggcaaaatttggcctttttgtgtgctgtgagcttgccacaaactgtactgtatctgag	649

Qy	1728	ttaacccgtaaacatcgcaatctgaagcaatcgaatctctgtcctgtggcggggtagtgtgag	1787
Db	650	tttatcamaacatmtccccaagccatgggtgtatctctgtgttgcctgggtttgtgtacg	709
Qy	1788	cgtaacggtgaaggtacgactctctatcagtatgaagaaagcacttaacgatatga---	1843
Db	710	aggaacccggtaaagcgaatgtttatccatcgaataatgatgtgaatctggtttatccaattta	769
Qy	1844	-----taagtgtctctgttcttggttcaaatgaacagagcgacccaagttaac	1889
Db	770	aaagatccacactctaaagtagcgcgtgatatatgatgcaaatatgaataaacacactgtgct	829
Qy	1890	cglttaacgttggcattaaacagccttaactatgcyggaanaatcccgatgaaggtcgt	1949
Db	830	cgtgcctccgggtacgtctgactgtggctgactgtgctgatacttccagagaccamaagat	889
Qy	1950	gagtctctattctcgtgtgtaa	1972
Db	890	caagatgtactgcctattattatga	912

RESULT	12
AL536748	
LOCUS	931 bp mRNA EST 13-FEB-2001
DEFINITION	AL536748 LTI_FLO13.Fbrn1 Homo sapiens cDNA clone CSDDF038YK02 5 prime, mRNA sequence.
ACCESSION	AL536748
VERSION	AL536748.1 GI:12800241
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eumaxyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 931)
TITLE	Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.
JOURNAL	Full-length cDNA libraries and normalization unpublished (2001)
COMMENT	Contact: Genoscope

Genoscope - Centre National de Sequencage
Bp 191 91006 Evry cedex - France
Email: segrete@genoscope.cns.fr, Web : www.genoscope.cns.fr
Location/Qualifiers
I . 931

```

SOURCE
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1="CSDF038K02"
/clone_1b="Lrl_FU013_FPrn1"
/dev_stage="pooled tissue from post conception fetuses (20
week, 24 week and 26 week)"
/lab_host="DH10B"
/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by Life
Technologies. Contact : Feng Liang life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT
215 a 211 c 269 g 234 t 2 others

```

	Query Match	Similarity	9.3%	Score 193.8	DB 106	Length 931
	Best Local	Similarity	55.2%	Pred No. 4.9e-34		
	Matches 390	Conservative	1	Mismatches 298	Indels 18	Gaps 1
Qy	1284	gcgcgtatttgcgcttgcaatcccaagaagtcagatcccaagaatataatgctccttaaat				1343
Db	207	gtgcagctccacagcttgcagatgagcagctccaccccaattcttaattccctggaaagtcgcaagcc				266

Oy	1344	gttgaacaggttttagtacttgaattcaacaacattaggttggttgatgtgcgtg	1403
Db	267	AGGGAGACCCAGACTGGTTTGGAGGTGGCCAGCACTTTGGGTGAGACACAGTAAGACT	326
Oy	1404	atcgcaattgcatcactgatgatthaacocggttttaagcgttaacaacatlaacnaac	1463
Db	327	ATTGCTATTGATGTGTAACAGAGCGTTGGTTAGAGCCAGAAAGTACTGGATTGCGTGA	386
Oy	1464	ccaatttcgtccagtyggaaacgaacaattgggtcgtatcatgaacgatttgggtgaa	1523
Db	387	CCATCAAAATTCCTGTGGTCTCTGAGACTTTGGGCAAAATCATGAATGATTTGGAGAA	446
Oy	1524	ccatcgtatgagaagtgtaaatcsgtctgagaagaaatggtctatccaccgtgcgcca	1583
Db	447	CCATTATATANAAGAGTCCCATCAAAACCAACATTTGCTCCCATTCATGTCGAGGCT	506
Oy	1584	ccaagttatgaagaacaactlaacagctactgaacttttagaaacgggaataaagttac	1643
Db	507	CCAGATTCATGGAATGACTGTGTGACGAGAAATTCGTGTACGTATCAAGGTTGTC	566
Oy	1644	gaacttagttgttcggttttggaagggttaagtaggtttatccggtggtgcggtgctc	1703
Db	567	GATCTGTAGCTCCCTATGAGCAAGGGGTGCANAATTGKCTTTTGGTGGTGGCGTGA	626
Oy	1704	ggtaaacocgtcaatatgatgaatlaatccgtlaaacctgcgaattgagcactcagttac	1763
Db	627	GGCAAGACTGTACTGATCATGATGAGGATTATCAACAATGTGCCAAAGCCCATGTCGTAC	686
Oy	1764	tctgtcttgcggggtagttagtcgttaacgctgagaaagtaaacacttatacagaagt	1823
Db	687	TCTGTGTTTCTGCTGTTGGTGTGAGAGACCCGTGAAGGCAATGAATTTTATCCATGAATG	746
Oy	1824	aaagacactlaacgctataga-----taagtgctcttgattatggt	1865
Db	747	ATTGAATCTGGTGTATCAACTTAAAGATGCCAACCCTCAAGGTAGCGGTGATATGCT	806
Oy	1866	caaatgaacgagccacacaggtaaacgctttagctgtgacatlaacaggtcttaactgyc	1925
Db	807	CAAAATGATTAACACCACTGGTGTCTGTCGCCCGGGTAGCTGTGACTGGGTGACTGTGCT	866
Oy	1926	gaaaaattccgtatgaaggttcgtgattgtcttattcttcgttgataa	1972
Db	867	GAATCTTTCAGAGACCAAGAGGCAAGATGTACTGCTATTATATGA	913

RESULT	13
AL516904	
LOCUS	AL516904 LITL_NF1D101.NBC1 Homo sapiens cDNA clone CSDDA007G13 5
DEFINITION	prlme, mRNA sequence.
ACCESSION	AL516904
VERSION	AL516904.1 GI:12780397
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 985)
JOURNAL	Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
COMMENT	Full-length cDNA libraries and normalization
CONTACT:	Unpublished (2001)
Genoscope	Contact: Genoscope

FEATURES	source	Location/Qualifiers
	1. .985	
/organism="Homo sapiens"		
/db_xref="taxon:9606"		
/clone_1fb="CSDDA007Yc13"		
/clone_1fb="LTI-NFL011.NBC1"		
/sex="male"		
/tissue_type="neuroblastoma cells"		

```

/lab.host="DH10B"
/note="Organ: brain; Vector: pCMWSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMWSPORT 6
vector. Library is not normalized, but is the control for
the normalized libraries. Library was constructed by Life
Technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
Maryland 20850, USA Fax : (1) 301 610 8371 Email :
liang@lifetech.com URL :
http://fulllength.invitrogen.com"
223 c 287 g 249 t 3 others

```

BASE COUNT	223 a	223 c	287 g	249 t	3 others
ORIGIN					

Query Match	9.38;	Score	183.4;	DB	105;	Length	985;
Best Local Similarity	55.08;	Pred. No.	6.1e-34;				
Matches	389;	Conservative	2;	Mismatches	298;	Indels	18;
				Gaps			1;

OY	1284	ggcggttttgacgttaattccccaagaagtgacgccaagaagtatatgatgccttaatt	1343
Db	207	GTGAGACTCCAGTTTATGATAGGAGACTACCAACCAATTTCTAAATGCCCTGGAACTGCAGAGCC	266
OY	1344	gttgaacacaggttttagtacttgaagtccaacaacaattagtgtygtgtgtatgcgtgt	14030
Db	267	AGGGAGACCCAGACTGGTGTGTGGAGGTGGCCCGACAGCAATTTGGTGAGAGCACAGTAAGAGACT	326
OY	1404	atgcgaatgsgatcalctgtatgtatataaacgcggttttaagtgatacaaatatcgaataac	14633
Db	327	ATTGCTATGGATGGTACAGAGAGGCTTGTTTAGAGGCGCAGAAAGTACTGGATTCGTGTGCA	386
OY	1464	ccaatttccttccagcttgaggaaagaacaatttgggtcgtatcatatgaacgtatttggttaa	15233
Db	387	CCAATCAAAATTCCTGTGTGCTCTGAGACTTTGGCGAGAAATCATGATATGTCTTGGAGAA	446
OY	1524	ccaatcgatgagcaagtgtaaatccggtgtcgaagaagaaatgtgtctattccacccgtgcgcca	15833
Db	447	CTATTGTATTAAGAAAGATGCCATCAACCAAAACCAATTTGCTCCATCATGCTGAGGCT	506
OY	1584	ccaagttatgaagaacaatctaacagctactgaactttagaaccgggaattaagaatttc	16433
Db	507	CCAAGTTCATGGAAATATGATGTGTGAGCAGAGAAATTCGTGTGATCTGTGATCAAGTTGTTC	566
OY	1644	gaacttagttgtccgttttgcgaaagggtttaaagttatttcgtgtgtgtgcggtgtc	17030
Db	567	GATCTGCTACCTCCCATATGCGCAAGGGGTGCAAAATTTGGGCTTTTGGGTGCTGAGACT	626
OY	1704	gttaaacgcgtcaatatgtatgaaattatccgtlaaacctgcaatttgaacatcagaattac	17633
Db	627	GGCAGAGCTACTGATCATGATGAGGATTATATCAACATATGTGCCAAAGCCCATGTGTGTTAC	686
OY	1764	tctgtcttgcgggggtlaagtgagcgtlaacgcgttgaagtaacgacttctatcatatgagtg	18233
Db	687	TCTGTGTTTCTGTGTGTGTGTGAGAGACCCGTGAAGCAATGATTTTATACATGAATG	746
OY	1824	aaagaccctacagctattaga-----taaagtgtcctctgtttatggt	18655
Db	747	ATTGTAATCTGTGTTATCAACTTAAAGAATGCCACCTCTAAAGTATGGCCTGTGATATGTG	806
OY	1866	caaatgaacagcagccacaggtlaacggttctttagctgtgagcattacaagcgttacaatgagtc	19255
Db	807	CAATATGAATGAACACACCTGCTGCTGTCGCCGGATGAGCTGTACATCGGGCTGACTGTGCT	866
OY	1926	gaaataatccgtgtatgaagtgctgtatgcttctatcttcctgtgataa	1972
Db	867	GAATATCTAGAGACCAAGAAAGTCTCAAGTATGATCTATTTATTTGA	913

prime, mRNA sequence.

VERSION AL518879.1 GI:12782372

SOURCE human.

Eukary

REFERENCE

NO.	TITLE	EXT./PAGES
1	Full-1	

COMMENT Contact

BP 191

FEATURES

source

BASE CONT 22

ORIGIN

Query Match

Best Local Similarity

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

—

[illegible]

100

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843

1004 CCaagLLa

DD : JLS CCAAGG11

	Query Match	Similarity	9.3%;	Score	183.4;	DB	105;	Length	985;
	Best Local	Similarity	5.0%;	Pred.	No. 6.1e-34;				
	Matches	389;	Conservative	2;	Mismatches	296;	Indels	18;	Gaps
OY	1284	gcggttattgaagcttgaattccccaagaatgagtaaccaagaatatalatgatgccttaaat	1343						
Db	213	GTGGAGCTGCATTTGATGAGGGATACCAACCAATTCTAAATGCCCTGGAAATGGCAAGC	272						
OY	1344	gttgaacaacggtttgttacttgaagtccaacaacattagtggtggtggttagttgcgct	1403						
Db	273	AGGGAGCCAGACTGGTTTTGGAGGTGGCCAGCAATTTGGGTGAAGCACAATAAGACT	332						
OY	1404	atcgcgaatgaggaatcatctgattgatttaaacgcggtttaagtcgttaacaatatcgataac	1463						
Db	333	ATTGCTAATGAGATGATGACGAAGAAGCTTGGTTAAGGCCAGAAAGTACTGGATTCTGGTGA	392						
OY	1464	ccaattctcttccagttgagggaagaacaacattgggtcgtatcatagaacgtattggtgaa	1523						
Db	393	CCAAATCAAAATTTCTGTTGGTCTTAGACTTTTGGGACAGATCATATAAATGTMAATTGGAGGA	452						
OY	1524	cgaatcgatgagcaagttgaaatcggtgcagaagaagaattggtctatccaacgttcgcga	1583						
Db	453	CCTATTGATTGAAGAAGGCTCCCATCAAAACCAACAATTGCTCCCATCATGCTAGGCT	512						
OY	1584	ccaattatgaagaacaatctaacagtaactcgaactttagaacccggaattaagaattac	1643						
Db	513	CCAGAGTTCATGAGGAATAGATTGAGCAGCAAGAAATTCGTGTCACGTGATCAAGCTTGTGC	572						
OY	1644	gaactaagtttgttcggttcggaagaagggtgaataagttatatttcgtgtgtgcggtgttc	1703						
Db	573	GATCTGTAAGCTCCTATTAGCCCAAGGGGTGGCAAAATATGGGCTTTTGGGTGGTCTGGAAGTA	632						
OY	1704	ggtaaacccgtcaatatgatgtaattaatcgttaacatcgaacattgagcaactcagttac	1763						

```

Db 633 GGCAGAGCTGCTACATGATCATGATGATATCAACATGTCGCCAAGGCCCATGTGTATAC 692
      || || || || || || || || || || || || || || || || || || || || || ||
Qy 1764 tctgtcttgcgggggtaagtgagtcgtaagcgtgaagtgtaacgactctctatgaatg 1823
      || || || || || || || || || || || || || || || || || || || || || ||
Db 693 TCTGTGTTGCTGCTGCTGTTGGTGAGAGACCCCGTGAAGCATGTTTATACCATGAATG 752
      || || || || || || || || || || || || || || || || || || || || || ||
Qy 1824 aaagactctaacgtattaga-----taagtgctctctgtttatcgt 1865
      || || || || || || || || || || || || || || || || || || || || || ||
Db 753 ATTGAATCTGCTGTTATCACTTAAAGATGCCACCTTAAGTAGCGGTATATGCT 812
      || || || || || || || || || || || || || || || || || || || || || ||
Qy 1866 caaatgaacgagccacacaggttaacggttgcgttgatgaacaggcttaactatg 1925
      || || || || || || || || || || || || || || || || || || || || || ||
Db 813 CAATGATGATGACACCTGCGTCTGCGCGGTAGCTCTGACTGGAGCTGACTGTGCT 872
      || || || || || || || || || || || || || || || || || || || || || ||
Qy 1926 gaaataatccgtgatgaagtcggtgatgtcttattcttcgttgtaa 1972
      || || || || || || || || || || || || || || || || || || || || || ||
Db 873 GAATACTTCAGAGACCAAGATGCAAGATGCTGCTATTATTATGA 919
      || || || || || || || || || || || || || || || || || || || || || ||

```

```

RESULT 15
LOCUS AL514536 1047 bp mRNA EST 13-FEB-2001
DEFINITION AL514536 LTR_NFL006_PL2 Homo sapiens CDNA clone CL08B003ZC04 5
ACCESSION prime, mRNA sequence.
VERSION AL514536.1 GI:12778030
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1047)
AUTHORS I., W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length CDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Source
1..1047
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CL08B003ZC04"
/clone_lib="LTR_NFL006_PL2"
/lissuse_type="placenta"
/note="Vector: pCMVSPORT 6. Site_1: NotI. 1st strand CDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded CDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 239 a 237 c 302 g 263 t 6 others
ORIGIN

```

```

Query Match 9.3%; Score 183.4; DB 105; Length 1047;
Best Local Similarity 55.0%; Pred. No. 6.2e-34;
Matches 389; Conservative 2; Mismatches 298; Indels 18; Gaps 1;

```

```

Qy 1284 gcggtatagcgttgatccacacagatgcagtaacaaagtatatgatccttaaat 1343
      || || || || || || || || || || || || || || || || || || || || || ||
Db 207 GTGGACGTCAGTTGAGTAGGAGGACTACACCAATTCCTAATGCCCTGGAAGTGCAGGC 266
      || || || || || || || || || || || || || || || || || || || || || ||
Qy 1344 gttgaacaggttgtagtacttaagtcaacacaaattaggtggtggtgtagtcgtg 1403
      || || || || || || || || || || || || || || || || || || || || || ||
Db 267 AGGAGACCAAGCTGCTGTTGGAGGYSGCCAGCATTTGGGTGAGAGACAGTAAGGACT 326
      || || || || || || || || || || || || || || || || || || || || || ||

```

```

Qy 1404 atcgcaatggatcatcctgtgtgattaaacgcggtttaaagcgttaacaaatacgaataac 1463
      || || || || || || || || || || || || || || || || || || || || || ||
Db 327 ATTGCTATGATGGATGATACAGAGGCTTGCTTAGAGGCCAGAAAGTACTGAGTTCTGGTGA 386
      || || || || || || || || || || || || || || || || || || || || || ||
Qy 1464 ccaatttcgttccagtgaggaaacacattggtcgtatcactgaacgtaattggtgaa 1523
      || || || || || || || || || || || || || || || || || || || || || ||
Db 387 CCATTCAAATTCCTGTTGGTCCCGAGACCTTGGCGAGAAATCATGATGATGATGAGAA 446
      || || || || || || || || || || || || || || || || || || || || || ||
Qy 1524 ccaatcgatgagcaaggatgaatcgtgtgcagagagaattggtctatcaacggtgc 1583
      || || || || || || || || || || || || || || || || || || || || || ||
Db 447 CCTATGATGAAAGAGGTCCCATCAAAACCAAAATTCCTCCATTCATGCTGAGAGGT 506
      || || || || || || || || || || || || || || || || || || || || || ||
Qy 1584 ccaagttatgaagacacatcacaagtaacgtaacttttagaagaggaaattaaagtatc 1643
      || || || || || || || || || || || || || || || || || || || || || ||
Db 507 CCAGAGTTCATGTGAATGAGGTTGACGACGAGAAATTCGTGAGCTGATCAAGGTTGTC 566
      || || || || || || || || || || || || || || || || || || || || || ||
Qy 1644 gacttagttgtccgttgcgaaagggttaaagtagttatcgttggtgcggtgctc 1703
      || || || || || || || || || || || || || || || || || || || || || ||
Db 567 GATCTGCTAGCTCCCTATGCCAAGGTGCGCAAAATTTGGGCTTTTGGTGGTGTGAGATT 626
      || || || || || || || || || || || || || || || || || || || || || ||
Qy 1704 ggtaaacgcgtcaatataatgaatgaatcgcgttaacatcgcaattgagcaactca 1763
      || || || || || || || || || || || || || || || || || || || || || ||
Db 627 GCGAAGACTGTTCTGATCATGAGATGATCAACAATGTCGCCAAGGCCCATGCTGTATAC 686
      || || || || || || || || || || || || || || || || || || || || || ||
Qy 1764 tctgtcttgcgggggtaagtgagcgtgaacggtgaaggttaacgactctatcatgagatg 1823
      || || || || || || || || || || || || || || || || || || || || || ||
Db 687 TCTGTGTTGCTGCTGTTGGTGAGAGACCCGTGAAGGCAATTTATACATGAATG 746
      || || || || || || || || || || || || || || || || || || || || || ||
Qy 1824 aaagactctaacgtattaga-----taagtgctctctgtttatcgt 1865
      || || || || || || || || || || || || || || || || || || || || || ||
Db 747 ATTGAATCTGCTGTTATCACTTAAAGATGCCACCTTAAGTAGCGGTATATGCT 806
      || || || || || || || || || || || || || || || || || || || || || ||
Qy 1866 caaatgaacgagccacacaggttaacggttgcgttgatgaacaggcttaactatg 1925
      || || || || || || || || || || || || || || || || || || || || || ||
Db 807 CAATGATGATGACACCTGCGTCTGCGCGGTAGCTCTGACTGGAGCTGACTGTGCT 866
      || || || || || || || || || || || || || || || || || || || || || ||
Qy 1926 gaaataatccgtgatgaagtcggtgatgtcttattcttcgttgtaa 1972
      || || || || || || || || || || || || || || || || || || || || || ||
Db 867 GAATACTTCAGAGACCAAGATGCAAGATGCTGCTATTATTATGA 913
      || || || || || || || || || || || || || || || || || || || || || ||

```

Search completed: October 28, 2001, 16:44:14
Job time: 6333 sec

10/31/01 13:40:41

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2001, 16:35:36 ; Search time 74.11 seconds
(without alignments)
5037.398 Million cell updates/sec

Title: US-09-545-199C-3

Perfect score: 1972
Sequence: 1 agcgaggccattgctcagc.....gtctattcttcgtgataa 1972

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 324599 seqs, 94655562 residues

Word size : 25

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCITUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	2.2	1374	2	US-08-743-637B-187 Sequence 187, App

ALIGNMENTS

RESULT 1
US-08-743-637B-187
Sequence 187, Application US/08743637B
Patent No. 5694066

GENERAL INFORMATION:

APPLICANT: BERGERON, Michel G.
APPLICANT: PICARD, Francois J.
APPLICANT: OUELLETTE, Marc
APPLICANT: ROY, Paul H.
TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA
TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND
IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED
TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS ...
NUMBER OF SEQUENCES: 273
CORRESPONDENCE ADDRESS:
ADDRESSEE: QUARLES & BRADY
STREET: 411 EAST WISCONSIN AVENUE
CITY: MILWAUKEE
STATE: WISCONSIN
COUNTRY: USA
ZIP: 53202-4497

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/743,637B
FILING DATE: 04-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/526,840
FILING DATE: 11-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: BAKER, Jean C.
REGISTRATION NUMBER: 35,433
REFERENCE/DOCKET NUMBER: 850586.90012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (414) 277-5000
TELEFAX: (414) 277-5591
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
LENGTH: 1374 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Haemophilus influenzae
US-08-743-637B-187

Query Match 2.2%, Score 44, DB 2, Length 1374;
Best Local Similarity 100.0%, Pred. No. 2.1e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 607 GTTATGTCGAATGACGACGACGAGGTAACGTTTACGTCGT 650
OY 1857 gttaatgccaatgaacgagccaccaggttaacggttaacgtgt 1900
|||||
|||||

Search completed: October 28, 2001, 17:56:16
Job time: 4840 sec

